AFRL-HE-WP-TR-2002-0156

UNITED STATES AIR FORCE RESEARCH LABORATORY

COGNITIVE PROBE PROJECT: DEVELOPMENT OF A TESTBED FOR COLLECTING COGNITIVE MODEL PARAMETERIZATION AND VALIDATION DATA

Edward A. Martin

HUMAN EFFECTIVENESS DIRECTORATE
CREW SYSTEM INTERFACE DIVISION
WRIGHT-PATTERSON AFB OH 45433-7022

JULY 2002

INTERIM REPORT FOR THE PERIOD MAY 2001 TO OCTOBER 2001

Human Effectiveness Directorate
Crew System Interface Division
2255 H Street
Wright-Patterson AFB, OH 45433-7022

NOTICES

TECHNICAL REVIEW AND APPROVAL

AFRL-HE-WP-TR-2002-0156

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Instruction 40-402.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

MARIS M. VIKMANIS
Chief, Crew System Interface Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2002	Interim, May 2001 - October 2001

4. TITLE AND SUBTITLE
Cognitive Probe Project: Development of a Testbed for Collecting Cognitive Model Parameterization and Validation Data

5. FUNDING NUMBERS
C: F41624-98-C-6012
PE: 61102F
PR: 7184
TA: 09
WU: 05

6. AUTHOR(S)
Edward A. Martin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Science Applications International Corporation (SAIC)
4031 Colonel Glenn Highway
Beavercreek OH 45431

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory, Human Effectiveness Directorate
Crew System Interface Division
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7022

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-HE-WP-TR-2002-0156

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
The Air Force Research Laboratory's Human Performance Model Integration (HPMI) program is exploring the merit and feasibility of combining multiple human performance models possessing dissimilar architectures to create integrated representations of human behavior that address application-specific requirements for model fidelity while controlling cost. The first exploration of HPMI feasibility involves the integration of a task-network model of a strike fighter pilot with an Adaptive Control of Thought-Rational (ACT-R) cognitive model. The ACT-R cognitive model is to replace the strike fighter pilot model's implementation of shootlist management, which is a dynamic and cognitively-intensive target-prioritization task. The Cognitive Probe project described in this report developed a concept and a virtual simulation testbed for obtaining the context-specific parameterization data needed to populate the ACT-R shootlist management model. One observation based on the data collection was that the earlier validation of the initial implementation of the strike fighter pilot model, which compared mission outcomes and task performance of the model against that of humans, was not sufficiently sensitive to cognitive aspects to pick up differences in shootlist management search strategies. Cognitive testbeds such as this may be beneficial for defining or verifying effective tactics and interface configuration before developing human performance models.

14. SUBJECT TERMS
cognitive modeling, cognitive probe, human behavior representation, human performance modeling, task network modeling, pilot models, crew system design

15. NUMBER OF PAGES
75

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNL

Standard Form 298 (Rev. 2-89) (EG)
Prescribed by ANSI Std. Z39-18

This page left blank intentionally

PREFACE

This effort, "*Validation of a cognitive model of a complex military information-management task*" aka "*Cognitive Probe Project,*" was conducted under contract number F41624-98-C-6012 with the Crew Systems Development Branch, Crew System Interface Division, Human Effectiveness Directorate of the Air Force Research Laboratory (AFRL/HECI), Wright-Patterson Air Force Base, Ohio 45433-7022, for the period May 2001 to October 2001. Science Applications International Corporation (SAIC), 4031 Col Glenn Highway, Beavercreek, Ohio 45431-7753 was the contractor. Dr. Edward Martin (AFRL/HECI) was the Principal Investigator. This effort was funded by discretionary funding from the Office of the Chief Scientist, Human Effectiveness Directorate, and supported Work Unit 71840905, "Human Performance Model Integration (HPMI)."

ACKNOWLEDGEMENTS

The "Requirements and Planning Document for the Cognitive Probe Component of the Combat Automation Requirements Testbed (CART) HPMI Effort" developed by Mr. Jeff Doyal of SAIC defined the testbed and data collection requirements. Ms. Karen Gery, also of SAIC, completed the testbed, scheduled the subjects, collected the data, and provided the summarized data to AFRL/HECI. The author is indebted to Mr. Doyal, Ms. Gery, Mr. Bryan Brett of SAIC, Dr. Christian Lebiere of Carnegie Mellon University (CMU), and Mr. Eric Biefeld of CMU for their many valuable contributions to this Technical Report.

This page left blank intentionally

TABLE OF CONTENTS

LIST OF FIGURES

This page left blank intentionally

LIST OF TABLES

This page left blank intentionally

"Human behavior representation is critical for the military services as they expand their reliance on the outputs from models and simulations for their activities in management, decision making, and training." (Pew and Mavor, 1998, p. 8)

INTRODUCTION

The Human Performance Model Integration (HPMI) Program

In its *Modeling and Simulation (M&S) Master Plan*, the Defense Modeling and Simulation Office (DMSO) has identified the capability to robustly represent individual and group behaviors as a critical need (DoD 5000.59-P, 1995). In a study commissioned by DMSO, the National Research Council's (NRC) *Panel on Modeling Human Behavior and Command Decision Making: Representations for Military Simulations* reviewed a number of architectures and tools that support the representation of various aspects of human behavior (Pew and Mavor, 1998). The panel pointed out that the architectures reviewed can be viewed as useful, promising, and a good starting point – *but are only very early steps* (Pew and Mavor, 1998, p. 337). The panel went on to say, "It is not likely, even in the future, that any single architecture will address all modeling requirements" (ibid.), and made the following observation.

> "It might be thought that a modular approach, in which modules are selected within each architecture and are then imported for use within other architectures, is a sensible compromise. However, a choice of modular boundaries is a choice of model conception, and one leaders in the field are not yet ready to provide. Thus we recommend that the architectures pursued within the military focus initially on the promising approaches identified in Chapter 3. This is an especially important point because the time scale for architecture development and employment is quite long..." (Pew and Mavor, 1998, p. 338)

In essence, the panel found that – collectively – the architectures and tools reviewed "offer a foundation on which to build models that will be truly useful and practical for military simulations" (Pew and Mavor, 1998, p. 110), but further work is needed before settling on any architecture. The panel discussed a hybrid approach to better encompass human phenomena (Pew and Mavor, 1998, pp. 108-111). In the panel's opinion, the most fruitful hybrid approach

would be interfacing architectures via communication protocols – rather than reimplementing features of one architecture in another (ibid., p. 109).

The Air Force Research Laboratory's Human Effectiveness Directorate has initiated the HPMI Applied Research Program to explore the feasibility and utility of hybrid human performance representations incorporating models with dissimilar architectures via standardized communication protocols. Architectures and tools currently under consideration are those that have been identified and reviewed by the NRC Panel (Pew and Mavor, 1998, pp. 51-111), with applications to military crew system interface domains. The motivation driving the *integration* of human performance models is to exploit available and proven modeling technologies as a means of incrementally providing more realistic representations of operator behavior. This approach also enables the selection of existing modeling technologies to satisfy application-specific fidelity requirements while controlling the cost of developing the model.

Relationship of the Cognitive Probe Project to the HPMI Program

The Cognitive Probe work reported here was conducted as a Basic Research Project under the HPMI Applied Research Program. The principal objective of the Cognitive Probe Project was to develop techniques for collecting cognitive data needed to support the initial HPMI model integration project, and – later – to support validation of resulting cognitive models. The Cognitive Probe Project is described in detail later.

The Initial HPMI Model Integration Project

As a first step toward the HPMI goal, the HPMI program is investigating the practicality of integrating models built using the following:

- Adaptive Control of Thought - Rational (ACT-R) architecture
- Micro Saint-based network tools

Adaptive Control of Thought - Rational (ACT-R)

ACT-R is one of the leading cognitive architectures, and has been under development over the past two decades (Anderson and Lebiere, 1998; Pew and Mavor, 1998, pp. 54-59). ACT-R

arguably covers the broadest range of psychological phenomena of any existing cognitive architecture. It combines a classical symbolic system with a neural network-like subsymbolic system. At the symbolic level, ACT-R implements a production system theory that models the steps of cognition through a sequence of production rules that fire to coordinate retrieval of information from the environment and from memory. At the subsymbolic level, every step of cognition implements parallel pattern matching that is tuned statistically to the structure of the environment.

ACT-R is a detailed cognitive theory that has been validated by hundreds of psychology experiments (ibid.). It can be used to model a wide range of human cognition — from tasks as simple as memory retrieval (Anderson et al., 1998) and visual search (Anderson et al., 1997) to tasks as complex as learning physics (Salvucci and Anderson, 2001) and air traffic control (Lee and Anderson, 2001). In all domains, it is distinguished by the detail and fidelity with which it models human cognition.

Micro Saint-Based Network Tools

Micro Saint is a simulation language and collection of tools that enable the construction of task network models for predicting human performance in complex systems (Pew and Mavor, 1998, pp. 71-75). These task network models are relatively easy to build, and can be understood by the practicing engineer or computer scientist. Task network models can be created at selected levels of detail, and are thus compatible with constructive simulations across a range of levels of aggregation. Further, a task network model-based human performance model (HPM) of a pilot performing an operationally-realistic mission has recently been developed, validated, and is available for use by the HPMI program. This task network model is discussed below.

The Task Network Model (TNM)-based HPM of a Strike-Fighter Pilot

TNM Implementation

The Air Force Research Laboratory Human Effectiveness Directorate's Combat Automation Requirements Testbed (CART) Advanced Technology Development Program is conducting case studies with the objective of demonstrating the CART program's concepts and tools. The

recently completed first case study involved the development and validation of an HPM of a single-seat, fighter-aircraft pilot conducting a one-ship Time Critical Target (TCT) mission (Brett et al., 2002). This case study used the existing Air Force Simulation Analysis Facility (SIMAF)[1] simulation environment architecture depicted in Figure 1. This simulation environment included: (1) a virtual cockpit interface (the Fighter Requirements Evaluation Demonstrator or 'FRED' cockpit shown in Figure 2), (2) a mission modeling environment (the Joint Integrated Mission Model or 'JIMM'), and (3) a shared memory interface connecting FRED and JIMM. For this case study, a Joint Strike Fighter Pilot HPM was developed on a stand-alone Microsoft Windows® platform, and then integrated with the FRED-JIMM simulator via High Level Architecture (HLA) Runtime Infrastructure (RTI) as illustrated in Figure 1. Middleware was added to the FRED cockpit interface to allow the HPM to control the simulated aircraft systems in place of a human operator during constructive HPM simulation runs. HPM validation results and TCT mission scenario task descriptions are summarized below.

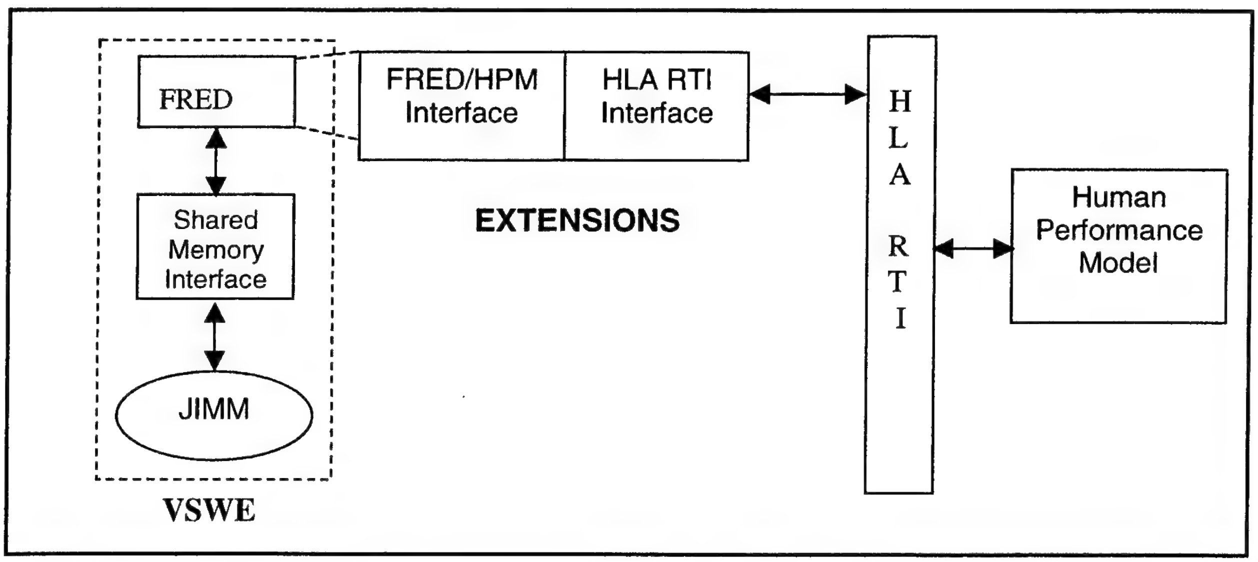


Figure 1. Depiction of the Integrated CART Case Study 1 Simulation Environment

<u>Strike-Fighter Pilot HPM Validation</u>

After developing the HPM and collecting constructive-simulation data with the HPM conducting the TCT mission in each of six different scenarios, virtual-simulation data were collected with

[1] SIMAF is located at Wright-Patterson Air Force Base, OH 45433-7505

eight pilots conducting the same six mission scenarios. These pilots interacted with the simulation through the original FRED cockpit interface. These data were then analyzed to determine how well the HPM predicted the pilots' performance. The HPM was found to have accounted for 61% of the variation of the behavior of pilots (Brett et al., 2002; Martin et al., 2001). Stated another way, the dependent measures from the constructive HPM simulation trials correlated with the data from the virtual human-pilot simulation trials at 0.78. As reported by Brett et al. (2002), this model did very well, on the whole, in predicting the pilots' performance.

In the following descriptions of the TCT scenario and shootlist management task, the 'pilot' tasks and procedures that are described had to be accomplished by either the actual pilots in a virtual simulation mode or by the HPM in a constructive simulation mode.

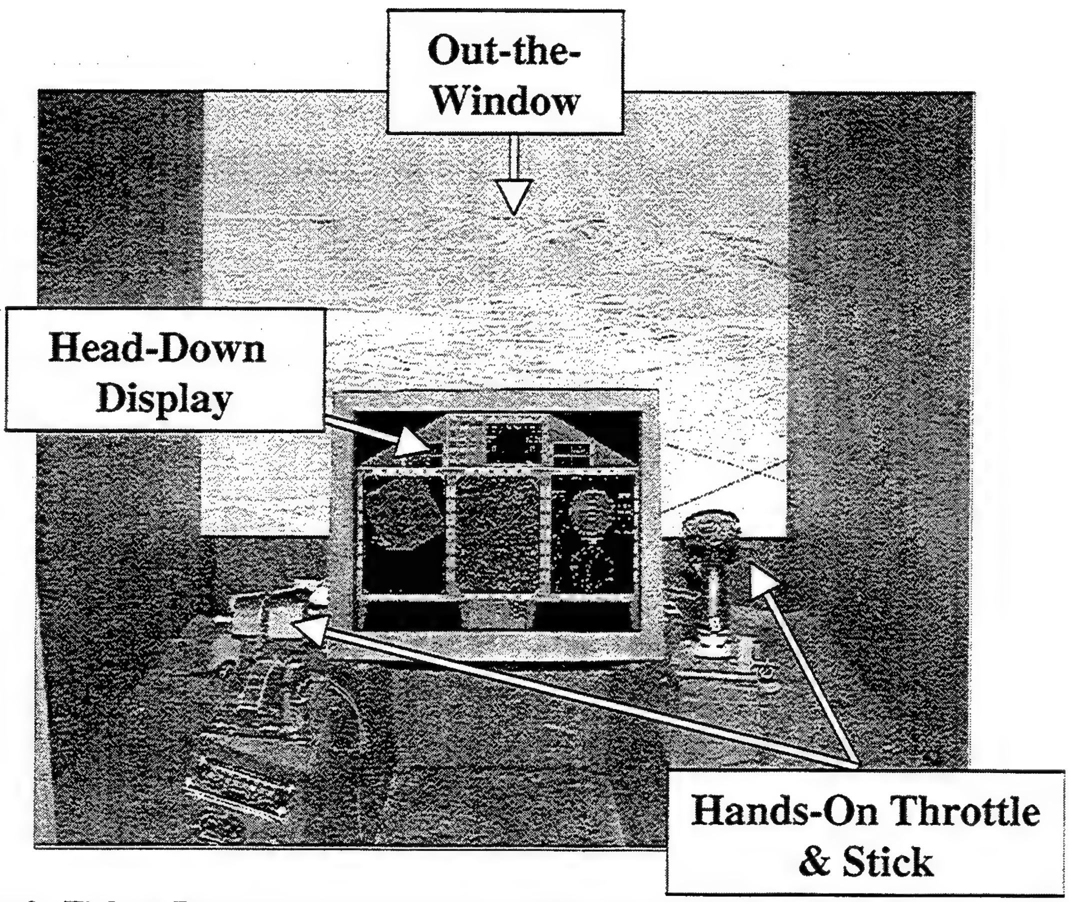


Figure 2. Fighter Requirements Evaluation Demonstrator (FRED) Cockpit and Displays.

Time Critical Target (TCT) Mission Scenario Description

The case study re-used TCT mission scenarios previously developed and used in Virtual Strike Warfare Environment (VSWE) exercises. TCTs are high-value, fleeting targets such as tactical ballistic missile launchers. The six TCT VSWE scenarios used were complex, highly dynamic, and operationally-realistic, and were drawn from earlier Air Force strike-fighter virtual simulations supporting weapon system trade studies. Figure 3 depicts the general form of the TCT scenarios. The scenario called for the pilot to employ multiple sensors to acquire and attack the mobile target. During ingress, the pilot was required to evade pop-up threats that launched surface-to-air missiles (SAM), and to subsequently recapture the ingress route and resume target acquisition. In addition, the pilot received and acted upon an in-flight intelligence update that provided a more accurate current location for the target. If the target was successfully acquired prior to arrival at the planned weapon release point, the pilot attacked it. Otherwise, the pilot was required to perform a manual re-planning activity in which a new route to re-fly the target area was developed. Once on this re-fly route, the pilot continued to attempt target acquisition and attack.

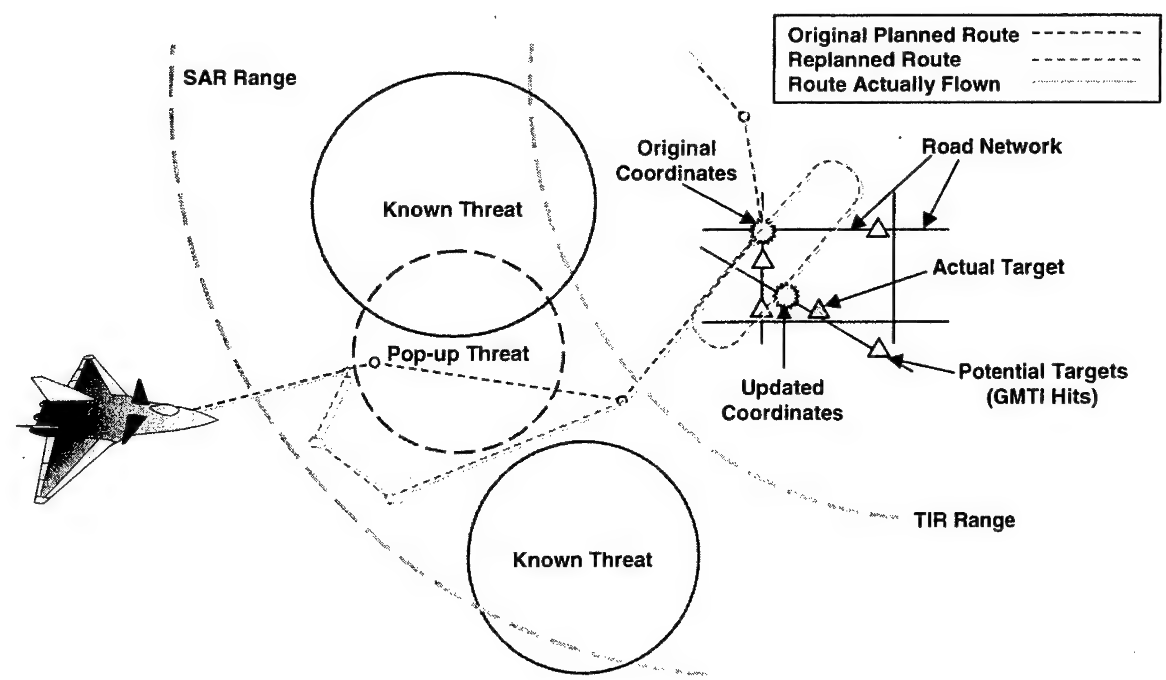


Figure 3. Illustration of Key TCT Mission Components

The Shootlist Management (SLM) Task

In the CART case study, the pilot's task was to search for a moving Scud Transporter / Erector / Launcher (TEL) using onboard sensors and to subsequently destroy it. The pilot had to fly the aircraft, navigate to the target area, evade threats, and acquire and attack the target. While enroute to the target area, a critical component of the pilot's sensor employment was the development and management of a shootlist (i.e., a prioritized target list).[2] The pilot initially detected targets using the radar's Ground Moving Target Indicator (GMTI), designated these detections to add them to the shootlist, and then attempted to identify shootlist items using a targeting infrared (TIR) sensor when within range (i.e., within 20 NM). Once objects were identified on the TIR display as something other than the TEL, the pilot could drop them from the shootlist. Moving and stationary objects viewed on the TIR display could also be detected and added to the shootlist. Items could be added or deleted from the shootlist as desired; however, the maximum number of shootlist items at any given time was nine.

To recapitulate, shootlist management (SLM) required the pilot to assess – on the fly -- the likelihood of sensor image components and GMTI hits coming from the TEL rather than from the other moving vehicles within the TEL's vicinity. The pilot had to prioritize detected targets within the shootlist, and then identify the shootlist items with the TIR before launching an attack or removing the item from the shootlist. This complex and dynamic information-processing task provided ample opportunity for human errors due to confusion, forgetfulness, or inappropriate prioritization of the shootlist. Unfortunately, as explained below, these cognitive aspects of human performance could not be well represented in a TNM.

[2] Although stored as a list in the mission computer, the shootlist does not appear as an actual list on the cockpit displays -- rather it is a number of icons on the tactical situation and radar displays that highlight the points / objects of interest that have been designated by the pilot as they were detected. The shootlist is built and modified dynamically over the course of the mission.

SLM Task Representation in a Strike-Fighter Pilot HPM

The *key cognitive components* associated with the SLM task discussed above are

- the ability to judge objects' range to the reference point, and to prioritize objects by this range, and
- the ability to remember which detected items have been identified, and to avoid taking multiple looks at the same object.

Original Implementation of the SLM Task

In the original CART strike-fighter pilot HPM, the strategy for determining which items to add to the shootlist and which shootlist items to examine next was based upon the criteria included in Table 1. Representation of the SLM task was modeled rather simplistically. It managed the mechanics of shootlist development, but did not represent underlying cognitive processes that could produce errors and other effects (e.g., the HPM always appropriately prioritized the targets, did not get confused, and did not forget). As such, the original HPM achieved perfect

Table 1. Shootlist Prioritization Criteria

Moving Status	Moving objects are given higher priority, as the scenario involves a mobile target that is likely to be on the move.
Range to the Reference Point†	Objects are prioritized based on their range from the reference point. Closer objects are given higher priority.
Time Since Failed Identification Attempt	Objects are temporarily assigned lower priority if the pilot attempts to identify them with the highest resolution sensor, and is unable to do so due to the aircraft's range from the object. Rather than dwelling on the object until it comes within identification range, the pilot will attempt to identify other objects on the list. The pilot will return to the particular object at a later time, when the aircraft is closer to it.

† Prior to the start of target acquisition activities, the shootlist contained only the 'Reference Point.' The 'Reference Point' was the latitude and longitude representing the best estimate of the target location. Since the target was a mover, the reference point position was stale by the time the pilot was trying to acquire the target.

accuracy with regard to the shootlist management criteria outlined in Table 1. For the initial HPMI model integration project, it was decided to model the SLM task using ACT-R as a way to include the human cognitive processing attributes in the HPM. The ACT-R model will implement the same basic object prioritization strategy as the original HPM, but will also include cognitive processes and effects that can potentially degrade 'pilot' performance.

New ACT-R Implementation of the SLM Task

For this first HPMI model integration project, the SLM component of the strike fighter pilot HPM will be removed from the existing software and recreated as an ACT-R model that will then be interfaced with the TNM via an external model call (EMC) from the TNM.[3] The ACT-R architecture will allow a more realistic representation of the cognitive components of the SLM task, allowing for memory decay and other cognitive phenomena to be encompassed. In order to develop an ACT-R based SLM model, cognitive performance data are required for model parameterization.

COGNITIVE PROBE PROJECT OBJECTIVES

The decision to implement the SLM task in an ACT-R model led to the requirement for an approach to capture key cognitive task data from live operators performing SLM. The Cognitive Probe Project was initiated with the following immediate objectives.

- Develop a concept for collecting human performance data to parameterize and validate a cognitive model of a complex, highly dynamic, operationally-realistic information-processing task.
- Develop a virtual simulator testbed to exercise the concept.
- Collect cognitive model parameterization data for the initial ACT-R model.

[3] EMC allows a task within the CART TNM to communicate with an external model using Microsoft's Component Object Model (COM) protocols. The EMC is set up by the modeler using a CART-developed Graphical User Interface (GUI).

COGNITIVE PROBE TESTBED DEVELOPMENT

A testbed was needed that would support immediate data collection for model parameterization and later data collection for model validation. To capture the necessary human-in-the-loop (HITL) data, a testbed was developed in which human subjects could fly a simulated strike mission against a TCT using the same cockpit interface and mission environment as used with the CART HPM. This Cognitive Probe Testbed was derived from a demonstration showpiece that had been originally developed for the CART program.[4] This involved modifying the CART demonstrator's FRED Advanced Virtual Cockpit interface simulation. As shown in Figure 4, the FRED Advanced Virtual Cockpit interface was modified through the addition of a 'fly box' in which was mounted two lever controllers (not currently used), eight pushbutton switches (only the two lower switches were used), and a BGSystems joystick. The joystick was configured with the thumb-actuated cursor controller and two four-position castle / coolie hat switches depicted in Figure 5, plus a trigger switch and a push button.

Advanced Virtual Cockpit (AVC) Sensor Displays

The three sensor displays shown in Figure 4 provided for acquisition and identification of the TCT. The left display panel is the real beam radar with a GMTI overlay,[5] the middle is the tactical situation display (TSD), and the right panel is the targeting infrared (TIR) display. Controls on the BGSystems Joystick, depicted in Figure 5, provided the means to slew the cursor, add (or remove) items to (or from) the shootlist and designate Next-to-Shoot (NTS), and select the Display of Interest (DOI). Figure 6 (which depicts a close-up of the computer screen in Figure 4) shows the AVC display symbology discussed below.

---

[4] The CART demonstrator was developed to exhibit the capabilities of the TNM-based strike fighter pilot model successfully prosecuting the TCT hunt. This demonstrator included the FRED-JIMM simulation environment illustrated in Figure 1 packaged into a two-processor Silicone Graphics Incorporated (SGI) Octane® computer.

[5] Unlike the CART case study, synthetic aperture radar (SAR) simulation was not available for the initial Cognitive Probe Testbed data collection.

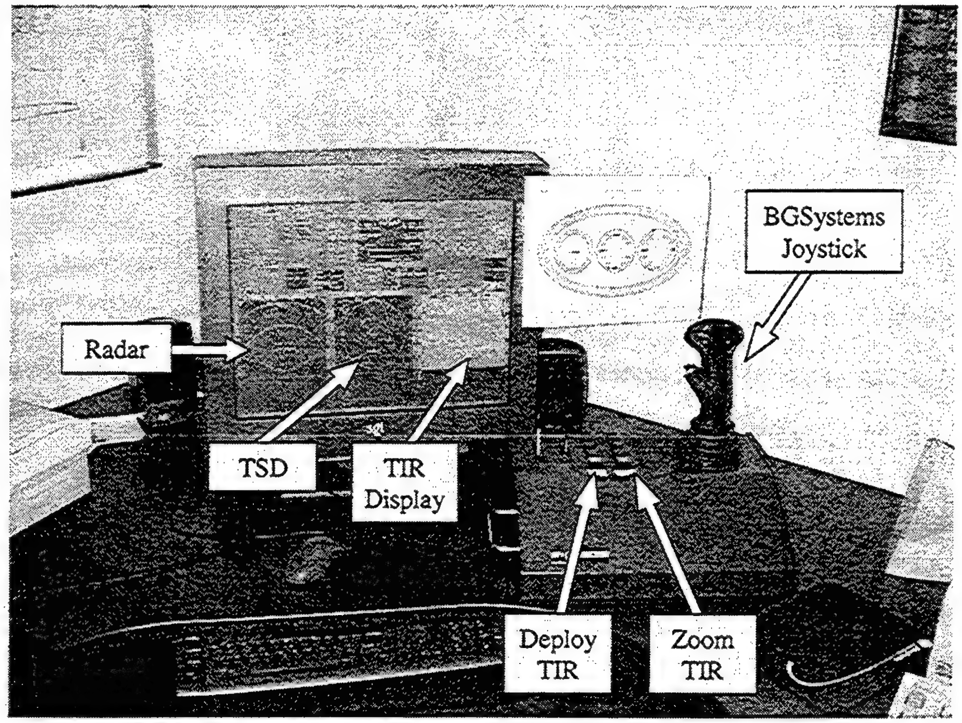


Figure 4. Cognitive Probe Testbed.

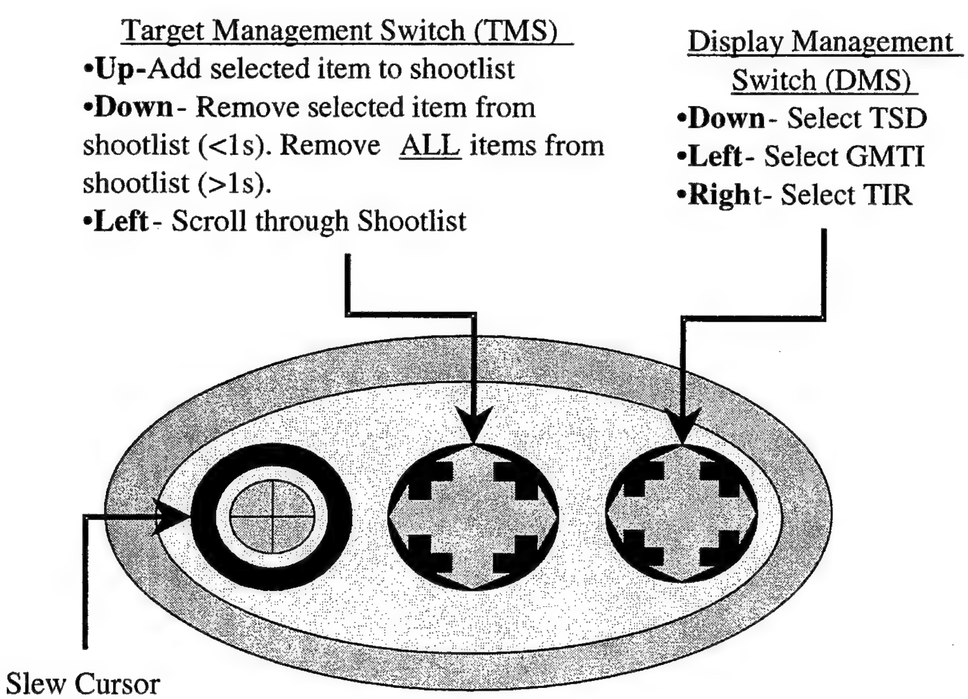


Figure 5. BGSystems Joystick Controls.

Radar Real Beam Ground Map Display

This display was used for target acquisition. Green 'eyebrows' (raw radar video) indicated a stationary object or a mobile object that the radar had not detected as moving.

The Ground Moving Target Indicator (GMTI) capability was automatically selected. Yellow dots were displayed at the location of any moving ground targets currently detected. Targets designated to the shootlist were denoted with a broken triangle. The current target of interest (TOI), also known as Next to Shoot (NTS), was denoted by a solid triangle.

Tactical Situation Display (TSD)

This display was used for target acquisition. The TSD displayed the aircraft's waypoints along with GMTI hits and other ground objects. The symbology used on the TSD is shown in Figure 6. Targets detected as ground vehicles are rendered as yellow circles wearing a square cap and other (i.e., stationary) ground targets are shown as yellow triangles. Targets designated to the shootlist are denoted with a broken triangle. The current target of interest (TOI), also known as Next to Shoot (NTS), is denoted by a solid triangle. An airfield (whose control tower position is used as the Reference Point in this project) is rendered as three intersecting runways.

Targeting Infrared (TIR)

This display was used for target identification. (The green highlighting in the corners in Figure 6 indicates that this was the current DOI.) The TIR sensor was deployed manually. This was accomplished by pressing the bottom left button on the fly panel. There were three different field-of-view (FOV) / zoom settings available for the TIR sensor (*wide*, *narrow*, and *2X narrow*). The bottom right button on the fly panel was used to cycle through these three levels. The current FOV or zoom level was indicated on the left side of the TIR display.

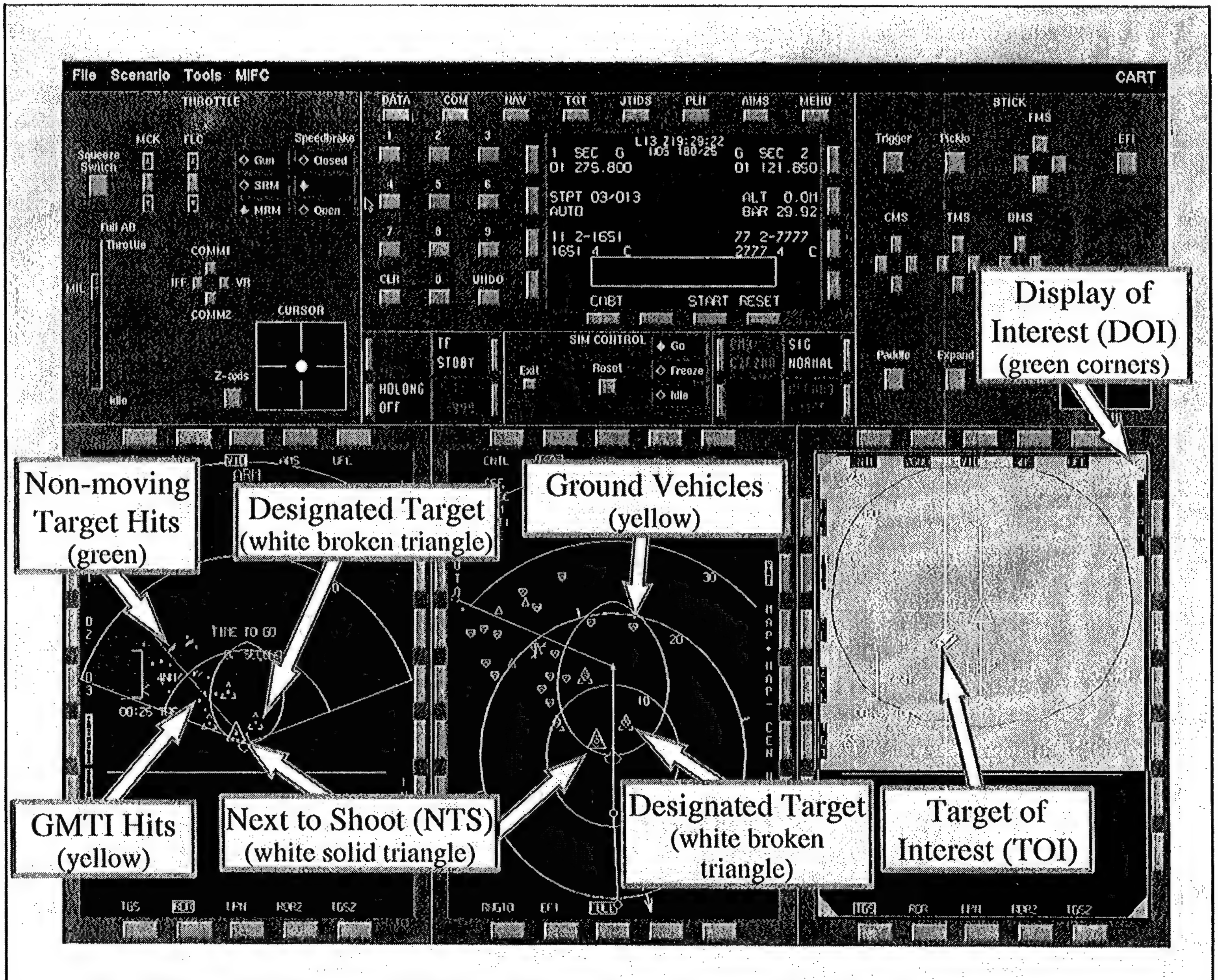


Figure 6. Advanced Virtual Cockpit (AVC) Interface Symbology.

Advanced Virtual Cockpit (AVC) Sensor Controls

Display Management Switch (DMS)

As shown in Figure 5, the DMS is the right-most button on the joystick, and it is used to control the Display of Interest (DOI). Moving the switch in any of three directions selects the corresponding display as the DOI. Moving the switch *right* selects the TIR, moving the switch *down* selects the TSD, and moving the switch *left* selects the radar with GMTI. The DOI is designated on the display by green highlighting in the corners.

Target Management Switch (TMS)

The TMS is the center button on the joystick (Figure 5). It permits the designation, rejection, and selection of target locations on multiple displays. The TMS functions described affect the TSD, radar, and TIR displays. Moving the TMS to the *forward* position designates the current cursor position. When a target location is designated it is added to the shootlist (the shootlist holds up to nine targets). Moving the TMS *forward* again makes that target the current target of interest (TOI) denoted by a solid triangle (also known as the Next To Shoot or NTS). Moving the TMS *left* steps through the shootlist. Subsequent movements continue cycling through the shootlist. Moving the TMS *down and releasing within one second* un-designates (or rejects) the current NTS-target. Rejecting a target that is not the NTS requires capturing the target by the cursor (cursor over the target) and moving the TMS *down and releasing within one second.* Moving the switch *down and hold for more than one second* clears the entire shootlist (removes all designations) .

Slew Cursor

The cursor controller provided for cursor slewing (including snap to object) on the radar, TIR, and TSD. Like in the original AVC user interface, this slewing was also used to change the zoom of the display.

Trigger

The joystick trigger was used to signal a positive identification of the desired target. Once a subject was satisfied that he or she had positively identified the TEL, the trigger was depressed. Trigger depression signaled the end of the trial.

Fly Panel

Eight buttons appear on the left of the joystick in Figure 4. Only the bottom two buttons were used.

- Deploy TIR: The lower-left, alternate-action button controlled deployment of the TIR. With the TIR retracted, pressing the button deployed the TIR and vice versa.
- Zoom TIR: The lower-right button controlled the field-of-view (FOV) or zoom level of the TIR. The three levels were *wide*, *narrow*, and *2X narrow*. Depressing the button moved through the levels from *wide* to *narrow* to *2X narrow*, and then back to *wide*.

METHOD

Subjects

For this initial data collection activity, subjects consisted of six SAIC engineers familiar with the simulation environment and acquisition task. Given the relatively simple, generic nature of the task (i.e., generic sensors, generic pilot-vehicle interface, no threats to deal with, and aircraft flight controlled by the autopilot), it was decided that the additional cost of bringing in Air Force pilots and training them on specifics of the mission and simulation would offer little additional benefit.[6] The subject ages ranged from 34 to 53, with a median age of 39.5 years. All were right-handed excepting one left-handed, 48 year old subject.

Procedure

Training, Practice, and Data Collection

The cognitive probe procedure was divided into three stages: (1) training, (2) practice, and (3) data collection. During training, the subjects were instructed on how to interact with the interface. Also, the details of the part-task TCT mission and the recommended search strategy were discussed. During training, subjects were led through each step. This was followed by a practice stage where subjects further familiarized themselves with the controls, and could ask questions to clarify any ambiguities. Subjects practiced until they could perform all critical tasks outlined in Table 2 without referring to any material such as the Figure 5-like 'cheat sheet' shown on the side of the monitor in Figure 4 (while this sheet was left in place during data collection, the researchers did not want subjects dependent upon it).

Once trained and practiced to the aforementioned level of proficiency, two data collection trials were conducted. The two data collection trials were conducted in the same order for all six subjects. No attempt was made to counterbalance for order effects in this initial study. Mission time for each of the two runs was approximately 15 minutes. Including training, practice, and between-scenario downtime, total time of participation for each subject was less than two hours.

Table 2. List of Critical Tasks Subjects Must Master Prior to Data Collection

CRITICAL TASKS CHECKLIST
Change Display of Interest (DOI)
Add item to shootlist on GMTI
Add item to shootlist on TIR
Remove an (1) item from shootlist
Remove all items from shootlist
Deploy TIR
Zoom in on GMTI
Zoom in on TIR
Inspect (scroll through) items on shootlist

Cognitive Probe Scenarios

Three scenarios were developed for the Cognitive Probe Testbed. One was used for practice trials, the other two were used for data collection. The terrain was based upon the Generic Composite Scenario (GCS) database used in prior VSWE exercises. In Appendix B, maps illustrate the initial lay down for each of these scenarios. The scenarios were similar in that they were threat-free, but the location of objects and waypoints varied. The data collection scenarios consisted of a three to four buildings and 20 to 21 movers. Each scenario included one TCT TEL that the subject attempted to locate and identify. The location of movers was unique for each scenario, and was dispersed throughout the gaming area to ensure GMTI hits did not overlap on the radar display. One airfield was included in each scenario; the control tower acted as the reference point.[7] The main gaming area extended out to a radius of approximately 20 miles.

---

[6] Unlike the CART case study scenarios, these initial training and data collection scenarios were free of threats and were flown solely under autopilot control.

[7] The control tower appeared in the TIR display, but not on the TSD or radar. The TSD display showed the location of the airport rendered as three intersecting runways.

The preplanned route established by the waypoints made one pass by the reference point, and contained waypoint-delimited legs appropriate for using the radar and TIR. The aircraft flew the preplanned route on autopilot; subjects did not control the simulated aircraft. Subject interaction was limited to managing the sensors, and locating and identifying the TEL.

Data Collection

Data from the simulation environment -- as well as subject-reported data -- were collected in an effort to determine the extent to which the subject could and did prioritize objects based on estimated range from a given reference point (Prioritization Data), and the extent to which the subject re-examined objects that had already been identified (Memory Data). In addition, subjective data regarding the pilot's SLM strategy were collected during post-session interviews (Subjective Strategy Data).

The three types of data identified for collection in the cognitive probe project are discussed below. Table 3 summarizes the nature of the prioritization and memory data collected.

Prioritization Data

Prioritization data measured how well the shootlist developed using the radar's GMTI capability corresponded to a normative shootlist based on the prioritization criteria of Table 1. From the start of a trial until a subject first deployed the TIR (it was also required that the reference point was within TIR range at this time), each item added to the shootlist was recorded. The list of objects retained on the shootlist at the time of first TIR deployment was referred to as the *Actual Global Prioritization List (GPL).* The maximum shootlist size could not exceed nine items. If an object was added to a full shootlist, the first item added would automatically drop off -- or be 'bumped' from the list (i.e., First-In, First-Out).

Additionally, from the onset of a trial until first deployment of the TIR, all objects detected and displayed by the GMTI -- as well as their range to the reference point at the first detection -- were recorded. These objects were prioritized and optimally sorted in accordance with the normative prioritization strategy (Table 1). These top nine prioritized objects formed the *Optimal GPL.*

Table 3. Prioritization and Memory Data Collection Summary

A pop-up menu on the display of interest was enabled each time that a shootlist item was removed from the given display (whether intentionally or by 'bump'). The simulation paused while the menu was active. The pop-up menu required the subject to select why an item was removed. Choices were *Removed-Identified, Removed-Unknown, Removed-Mistake, Bumped-Identified,* and *Bumped-Unknown.** The selection was then tagged to the ID of object removed.
A running list of all unique objects detected by GMTI was kept. Object data included ID, type, range to reference point, and simulation time of first detection (i.e., first appearance on a radar display). At the time of first TIR deployment within TIR range to the reference point, the optimal nine-item shootlist – using the prioritization criteria of Table 1 -- was created. This was the *Optimal Global Prioritization List.*
At the time of first TIR deployment within TIR range to the reference point, all items on the subject's shootlist were captured. Object data included ID, type, and range to the reference point. This was the *Actual Global Prioritization List.*
A time-stamped list of all shootlist add / remove events was kept. Object data included ID, type, range to reference point, event type (add or remove), removal rationale (from the pop-up menu selection) for removal events, and also the object's range from the aircraft starting position. (Although it did not directly impact ACT-R parameterization, this last item could provide insights into shootlist management *strategy*, and might prove useful for future data-collection efforts.) Also, the total number of unique objects added to the shootlist and the number of times a previously-removed object was re-added were captured. The maximum shootlist size during each trial was captured.

* These response options are fully defined in the *On-line Survey* section of Appendix C.

Note that the GPL data recorded were a snapshot of the subject's shootlist at the moment of first deployment of the TIR. The subjects further modified the shootlist after entering TIR range.

Memory Data

Memory data were recorded once the subject was within TIR range, and therefore could use the TIR to identify the objects. Memory data measured how often a subject re-assigned an object to the shootlist that had already been identified and removed -- thereby reducing acquisition efficiency. The concept for capturing memory data was to understand *why* a subject removed an item from the shootlist (e.g., whether it was not yet identified but had to be removed to make room for a higher priority detection, or whether it was removed after being identified as an object other than the TEL). Each time an object was added to or removed from the shootlist, the add or

remove event and the corresponding object ID were recorded. Each time an object was removed from the shootlist (whether it got 'bumped off' or was intentionally removed), the subject was asked to specify why the object was removed via a pop-up menu. This cognitive probe was necessary, as the subject's intention could not be divined from the simulator data.

Subjective Search Strategy Data

In addition to the objective data collected, a two-question questionnaire was administered to gain further insight into each subject's prioritization and search strategy. The purpose of the questionnaire was to help identify the subjects' strategies for building the shootlist. This was intended to provide insights regarding potential modifications necessary for the modeled search strategy. The questionnaire was given at the completion of the two data collection trials. Subjects were given no prior knowledge of the questions in order to preclude such knowledge from influencing their behavior. The questions were open-ended in nature:

- During the experiment, did you implement the shootlist strategy discussed in the instructions?
- Given what you experienced in the experiment, explain the strategy you would use to manage the shootlist to search and identify the TCT. (Consider your strategy in a real-world context that includes terrain, control of the aircraft, potential of threats, many more movers, more types of movers, etc.)

Subject responses to these two questions are included in Appendix D.

Data Analysis

The principal goal of the Cognitive Probe data analysis was to obtain meaningful data for better parameterizing the shootlist management function in the ACT-R model. The approach included comparisons of the subject's *Actual GPL* to the *Optimal GPL* for each trial in order to derive a mean prioritization error that could be applied to the prioritization strategy within ACT-R. Memory data were also analyzed in order to identify the probability that an object, once identified as being an object other than the TEL and consequently removed from the shootlist, would later be re-added to the shootlist.

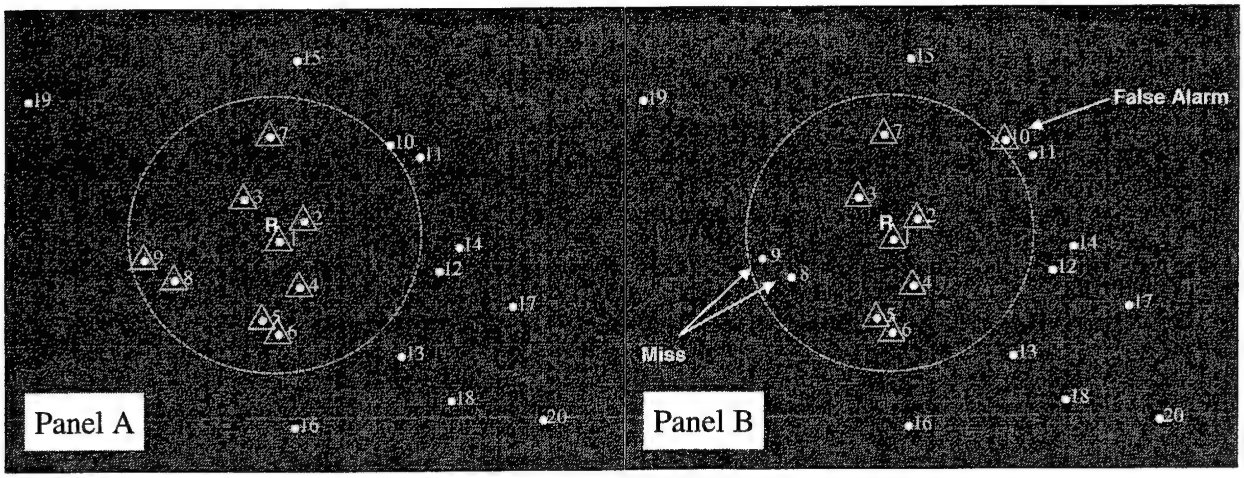


Figure 7. An illustration comparing an *Optimal GPL* with a notional *Actual GPL*. Panel A depicts an *Optimal GPL,* while Panel B portrays an *Actual GPL* containing *Miss* and *False Alarm* errors. R represents the Reference Point. The numbered dots represent GMTI hits. The dots enclosed in triangles represent those GMTI hits designated to the shootlist. The dashed circle encloses the nine GMTI hits closest to the Reference Point.

Prioritization Errors

Prioritization Error Categorization. Prioritization errors were categorized as being either errors of omission (i.e., *misses*) or commission (i.e., *false alarms*). Panel A of Figure 7 illustrates an 'Optimal GPL' constructed in accordance with the normative prioritization criteria of Table 1. Panel B of Figure 7 illustrates a notional 'Actual GPL' constructed with some deviations from Table 1's criteria leading to *miss* and *false alarm* errors. In Panel B, GMTI hit number 10 is seen to have been designated to the shootlist, when in fact there are nine other GMTI-identified moving vehicles closer to the Reference Point; this is termed a *false alarm*. It is also seen that GMTI hit numbers 8 and 9 have not been designated to the shootlist, when in fact they are closer to the Reference Point than GMTI hit number 10; these are termed *misses*. Note that -- by definition -- a *false alarm* cannot occur without a corresponding *miss*.

Prioritization Error Calculations. At the time the aircraft came within TIR range to the reference point and the TIR was first deployed, a list of all objects having been detected with the GMTI -- along with the objects' range-to-reference-point at first detection -- were recorded for each trial. All GMTI-detected objects were prioritized on the basis of their range to the reference

point, creating the *Optimal GPL*. An Optimal GPL is exemplified in Columns A-B-C in Table 4, which is a datasheet for one of the data collection trials.

Similarly, objects having been designated to the shootlist were also recorded at the time of first TIR deployment within TIR range to the reference point. These objects made up the *Actual GPL* for a given trial. Table 4 exemplifies Actual GPL data in Columns G through M. Summary data calculations in rows 26 through 44 are annotated to show the formula used.

A prioritization error was calculated for each item that appeared on an Actual GPL (indicated by a *Y* in column D of Table 4). The assumption underlying the prioritization error was that a subject's failure to prioritize shootlist items based on range-to-reference-point was due to the subject's inability to accurately and consistently estimate an object's range to the reference point. The following describes the approach used for calculating the magnitude of this error (refer to columns G through M of Table 4).

If the item on the Actual GPL also appeared on the Optimal GPL, that item was classified as a *hit* and the item's error was zero. Similarly, if an item not appearing on the Optimal GPL was also absent from the Actual GPL, it was considered a correct rejection (*CorrRej*) with an error of zero.

If an item on the Optimal GPL did not appear on the Actual GPL, that item was classified as a *miss*. Misses were classified as either *Type 1* or *Type 2*. If an item on the Actual GPL was further from the reference point than a missed item, it was inferred that the subject judged the corresponding object to be closer than the missed object. This was classified as a *Type 1 Miss*, and the associated error was calculated by subtracting the range to the reference point of the omitted Optimal GPL item from the range to the reference point of the furthest Actual GPL item. As an example, refer to Object ID 12's row in Table 4. Object ID 12, with a range to the reference point of 3.06 NM, did not appear on the subject's Actual GPL. The furthest item from the reference point that *did* appear on the subject's Actual GPL was Object ID 9 at a range of 9.82 NM to the reference point. In this case, we subtracted the range of the omitted item from the range of the furthest item on the Actual GPL to arrive at an estimated error associated with the omission of Object ID 12.

Table 4. Prioritization Error Datasheet for a Single Trial

	A	B	C	D	G	H	I	J	K	L	M	N	O	P
2		Optimal GPL			Actual GPL									
4	Object ID	Priority	Range to Ref (NM)	On SL?	Classification	Error Type	Range of Missed Items	Range Actual GPL Items	Miss Error (NM)	FA Error	Type 1 Miss Error			
5	21	1	2.15	Y	Hit			2.1537968						
6	12	2	3.06		Miss	Type 1 Miss	3.0632026		6.752306582		6.75230658			
7	6	3	3.51	Y	Hit			3.5088411						
8	22	4	4.68		Miss	Type 1 Miss	4.6815915		5.13391776		5.13391776			
9	18	5	4.69	Y	Hit			4.6899068						
10	26	6	5.08	Y	Hit			5.0836693						
11	23	7	5.27	Y	Hit			5.2666972						
12	24	8	6.38		Miss	Type 1 Miss	6.3753555		3.440153681		3.44015368			
13	10	9	7.01		Miss	Type 1 Miss	7.0073116		2.808197618		2.80819762			
14	11	10	7.21		CorrRej									
15	7	11	7.37		CorrRej									
16	15	12	7.82		CorrRej									
17	17	13	7.93	Y	FA	FA		7.9332552		4.870052507				
18	8	14	8.55		CorrRej									
19	25	15	8.89		CorrRej									
20	13	16	9.15		CorrRej									
21	9	17	9.82	Y	FA	FA		9.8155092		6.752306582				
22	19	18	10.43		CorrRej									
23	14	19	13.85		CorrRej									
24	20	20	13.91		CorrRej									

		K	L
25			
26	Number of Type 1 Miss Errors	4	"=COUNTIF(H5:H24,"Type 1 Miss")"
27	Number of Type 2 Miss Errors	0	"=COUNTIF(H5:H24,"Type 2 Miss")"
28	Number of False Alarm Errors	2	"=COUNTIF(H5:H24,"FA")"
29	Total Error Magnitude (NM)	18.135	"=SUM(M5:M24)"
30	Avg Error Magnitude (NM)	4.53	"=IF(K26>0,K29/K26,"")"
31	Number of Detected Items	20	"=B24"
32	Prob of Error	30.00%	"=SUM(K26:K28)/K31"
33	Avg Error Magnitude Across All Detections (NM)	0.91	"=SUM(M5:M24)/K31"
34	Shootlist Size	7	"=COUNTIF(D5:D24,"Y")"
35	Number of Hits	5	"=COUNTIF(G5:G24,"Hit")"
36	Number of Correct Rejections	9	"=COUNTIF(G5:G24,"CorrRej")"
37	Number of False Alarms	2	"=COUNTIF(G5:G24,"FA")"
38	Prob of False Alarm given Error	0.333	"=IF(SUM(K26:K28)>0,K37/SUM(K26:K28),"")"
39	Magnitude of False Alarm Error (NM)	11.622	"=SUM(L5:L24)"
40	Avg Magnitude of False Alarm Error	5.81	"=IF(K37>0,SUM(L5:L24)/K37,"")"
41	Number of Misses	4	"=SUM(K26,K27)"
42	Probability of Miss given Error	0.667	"=IF(SUM(K26:K28)>0,K41/SUM(K26:K28),"")"
43	Magnitude of Miss Error	18.135	"=SUM(K5:K24)"
44	Avg Magnitude of Miss Error	4.53	"=IF(K41>0,SUM(K5:K24)/K41,"")"
45			

	A	B		
46	Range of Furthest Actual GPL Item	9.815509229	Miss Type 1	If there are objects on the Actual GPL that are further from the reference point than the missed object, we can infer that the subject judged these objects to be closer than the missed object. As such, we can calculate an associated error by subtracting the range to the reference point of the omitted Optimal GPL item from the range to the reference point of *the* furthest Actual GPL item. Since miss magnitudes are calculated using the range of the furthest item from the reference point appearing on the Actual GPL, miss magnitudes will be larger in situations where false alarms are present as false alarms are by definition far from the reference point.
47	Range of Ninth Object on Optimal GPL	7.01	Miss Type 2	If there are no objects on the Actual GPL further from the reference point than a missed object, that missed object will not contribute to the error magnitude. The rationale for this is that we cannot assume that such an object was missed because the subject misjudged the objects distance from the reference point -- other factors, such as workload, might have caused a 'Miss Type 2' error.
48	Range of Missed Item Nearest to Ref	3.063202648	False Alarm	If an item on the Actual GPL does not appear on the Optimal GPL, that item will be considered a False Alarm. For every false alarm, there must, by definition, be a miss. Therefore, accumulating magnitude error for both the miss and the false alarm would be inappropriate 'double dipping.' False Alarm errors do not contribute to the error magnitude, only 'Miss Type 1' errors contribute.

Error for Optimal GPL Item (Object ID 12) Not Appearing in Actual GPL (Type 1 Miss)

9.81 NM – 3.06 NM = 6.75 NM Error Magnitude

Total Error Magnitude (NM) = Summation of Type 1 Miss Errors Across All Object IDs

6.75 + 5.13 + 3.44 + 2.81 = 18.13 NM Total Error Magnitude

$$\text{Average Error Magnitude (NM)} = \frac{\text{Total Error Magnitude (NM)}}{\text{Number of Type 1 Miss Errors}} = \frac{18.13}{4} = 4.53$$

If, on the other hand, there was no Actual GPL object further from the reference point than a missed Optimal GPL object, this was calssified as a *Type 2 Miss*. Type 2 Misses did not contribute to the error magnitude because it could not be inferred that the subject necessarily misjudged the object's distance from the reference point. Other factors, such as the subject's ability to handle the workload or short-term memory constraints, could contribute to Type 2 Misses.

If an item on the Actual GPL *did not* appear on the Optimal GPL, that item was considered a *False Alarm* (FA). For every False Alarm, there must – by definition – be a Type 1 Miss. Accumulating magnitude error for both the Type 1 Miss and the False Alarm would be counting the same error twice. Therefore, False Alarms did not contribute to error magnitude.

Memory Errors

Probability of Re-add data regarded the probability that an item, once identified as an object other than the TEL and subsequently removed from the shootlist, would later be re-added to the shootlist. When this occurred, we inferred that the subject did not remember / recognize that the object had already been identified, and intended to attempt identification again. This probability was calculated as follows:

$$\text{Probability of Re-adding} = \frac{\text{Number of Identified \& Removed Objects Later Appearing on List}}{\text{Number of Identified \& Removed Objects}}$$

In addition to the *probability* of re-adding an object, it was necessary to collect data regarding the *latency* associated with the re-add for the ACT-R model. This was done by subtracting the mission time at which an object was removed from the shootlist from the time it was re-added.

Re-add frequency data were also collected. Frequency of re-adds was calculated as the average number of re-adds per trial. This was calculated separately for each scenario.

RESULTS AND DISCUSSION

Prioritization Errors

The prioritization error results are summarized in Table 5 and Table 6 (see the *Data Analysis* section for definitions for the table entries). Data averages and standard deviations across the six subjects are presented in Table 5 for each data collection scenario. Table 6 presents the corresponding minimum, median, and maximum values.

Table 5. Prioritization Error Averages and Standard Deviations

	Scenario 1		Scenario 2	
	Average	SD	Average	SD
Shootlist Size	6.67	1.97	7.00	1.10
Number of Hits	4.67	1.97	4.67	3.01
Number of Correct Rejections	10.00	1.79	8.67	2.16
Number of Type 1 Misses	4.17	1.83	4.00	3.41
Number of Type 2 Misses	0.17	0.41	0.33	0.82
Number of False Alarms	2.00	1.79	2.33	2.16
Average Error Magnitude (NM)	6.48	5.10	5.31	3.15

Table 6. Prioritization Error Medians and Extremes

	Scenario 1			Scenario 2		
	Min	Median	Max	Min	Median	Max
Shootlist Size	3	7	9	5	7	8
Number of Hits	2	5	7	0	6	7
Number of Correct Rejections	7	11	12	6	10	11
Number of Type 1 Misses	2	4	7	0	3	9
Number of Type 2 Misses	0	0	1	0	0	2
Number of False Alarms	0	2	5	0	2	5
Average Error Magnitude (NM)	0.73	7.88	16.67	0.93	4.67	11.69

Shootlist size is seen to have ranged from three to nine, with a median value of seven. This is consistent with Miller's (1976) classical observation that human channel capacity limits the number of equally likely alternatives about which humans can make correct absolute judgments

to 7±2.[8] This suggests that if there is a future desire to expand the shootlist size beyond nine items, the operator will require additional aids such as cues regarding items intentionally removed from the list and priorities the operator may assign to shootlist items.

None of the subjects achieved nine correct hits in accordance with the prioritization criteria of Table 1. Each had two or more *miss* errors (the one subject who had no Type 1 Misses in Scenario 2 had two Type 2 Misses). This and experimenter observation both suggest that the subjects did not adhere to the normative search strategy as instructed (see Appendix C), despite the fact that they all reported that they did (see Appendix D). This should not be surprising, since other studies have shown that subjects will employ different strategies in the same task context, and will shift strategies as a way of optimizing human-system interaction while minimizing the cost of that interaction (Gray and Boehm-Davis, 2000). In the words of one distinguished researcher in the field of human cognition, "Subjects cannot maintain a consistent strategy although they try; subjects keep shifting strategy during the course of a threat classification trial."[9] Unfortunately, the Cognitive Probe researchers were not aware of this other work when they were collecting the data. It was the data from the Cognitive Probe Testbed that brought this insight to them. This also leads to an interesting observation regarding the CART case study, wherein the SLM strategy used in the Cognitive Probe Study was developed. The strike-fighter pilot HPM used in the CART study implemented the SLM task flawlessly in accordance with the prioritization criteria of Table 1. Yet the human operator performance against which the HPM was validated most likely did not apply the strategy of Table 1 consistently. Clearly the CART case study measures were not sufficiently sensitive to cognitive performance to show these differences, and perhaps lead the CART researchers to question the shootlist-search strategy they developed. This is one example in which a cognitively-oriented

[8] Miller referred to this as the *span of absolute judgment*. He also introduced the notion of *chunks* to distinguish this from the *span of immediate memory* (chunks are more loosely defined than bits, and can aggregate bits of information through the process of *recoding*). A summary is provide in Boff et al. (1986, pp. 41-6 to 41-8)

[9] Wayne Gray, Department of Psychology, George Mason University, personal communication 9 January 2002. The "threat classification trial" referred to is in many ways similar to the shootlist management task (Schoelles and Gray, 2000).

testbed could have played an important role in providing insight regarding cognitive performance in a system.[10]

In the previous two paragraphs are illustrations regarding how a cognitively-oriented testbed might have been valuable for defining or evaluating interface capabilities and effective tactics (or at least for providing decision makers insights regarding deficiencies), beyond its value in supporting HPM parameterization or validation. In the first instance, data from the testbed clearly show that the interface – as designed – will not support SLM for lists containing more than nine items (for most of the population, the interface will not even support nine entries). The second instance illustrates how such a testbed might foster an awareness that the tactics actually being employed are substantially different from what the designers and the operators think.

Despite the observation that the Cognitive Probe subjects did not consistently adhere to the normative search strategy, and were apparently unaware that they deviated from the search strategy that they were instructed to follow, their target acquisition performance was comparable to that of the CART case study subjects. The CART case study subjects correctly acquired the TEL 98% of the time (47 out of 48 trials). The Cognitive Probe subjects correctly acquired the TEL 92% of the time (11 out of 12 trials). Further, the CART HPM -- which followed the normative search strategy flawlessly -- correctly acquired the TEL 100% of the time (36 out of 36 trials), and had an overall correlation with the pilot-in-the-loop measures of 0.78 (Martin et al., 2001). As reported by Craig et al. (2002), it is suspected that the test scenarios were too simple (in terms of the number of potential targets to be examined) to demonstrate differences in shootlist management performance due to differences in cognitive behavior. It is expected that if the number of targets to be examined is increased to better stress cognitive shootlist management performance, that mission performance will degrade relative to performance predicted by the original CART case study HPM implementation. It is also anticipated that the performance predicted by the hybrid TNM with an ACT-R implementation of the SLM task will better represent that of humans performing the same task.

[10] Experience with the Cognitive Probe Testbed suggests that pop-up probes provide an effective means for capturing elements of cognition with minimal disruption.

Memory Errors

There were twelve instances of shootlist items being identified as an object other than the TEL, removed from the shootlist, and later re-designated to the shootlist. These all occurred in Scenario 1. There were no such instances in Scenario 2; in the judgment of the researchers this was due to the TEL being located on the near side of the reference point relative to the aircraft's initial position in Scenario 2 (refer to Appendix B). In fact, Scenario 1 was deliberately constructed such that the TEL would not appear in the nine top-priority GMTI hits. If the TEL were included in the top nine entries of the shootlist, it could be identified with no need for remove and add actions – and thus no memory data. Therefore, Scenario 1 was constructed in a way that provided the subjects an opportunity to forget.

The *re-add* instances in Scenario 1 involved an equal number of tanks and trucks. The highest number of *re-adds* involved the vehicles closest to the reference point – four *re-adds* for the closest, two for the second closest, one each for the next four closest, and then one each for two of the more distant vehicles. The number of *re-adds* varied across subjects. Two subjects re-added four vehicles, one subject re-added three, one subject re-added one, and two subjects re-added none. Only one subject re-added the same vehicle more than once; this subject re-added the vehicle closest to the reference point three times. The *re-add* data are summarized in Table 7 for Scenario 1; there is no table for Scenario 2 since there were no instances of *re-add* for that scenario.

Recall from the *Data Analysis* section that the *probability of re-adding* equated to the probability of forgetting that a particular vehicle had been previously identified, and removed from the shootlist. The *re-add frequency* indicates how often a vehicle might be re-added to the shootlist during the course of a trial. The *re-add latency* provides some insight into memory retention

Table 7. Memory Error Data for Scenario 1

	Average	SD	Min	Median	Max
Probability of Re-adding	0.23	0.26	0	0.17	0.67
Frequency of Re-adding	1.20	0.63	1	1	3
Re-add Latency (seconds)	43.45	35.95	4.08	34.70	124.34

times for SLM. While re-add latency does not equate to memory retention time per se (a subject would not necessarily re-add a vehicle immediately upon forgetting that it had been intentionally deleted), it does provide an upper-bound value.

THE WAY AHEAD

The next step is to use the data collected to parameterize an ACT-R model of the SLM task, and then conduct model runs to explore and benchmark the performance of the hybrid HPM.

The ACT-R model will execute the same normative search strategy (Table 1) as the CART case study model. However, ACT-R will account for cognitive processes and effects that better represent human cognitive limitations. In ACT-R, the matching of chunks in memory to production conditions is not a perfect process. Rather, chunks of the same type compete with chunks that only partially match the desired pattern. Retrieval of a chunk from memory is a stochastic process, and – as with a human – sometimes results in error. This Partial Matching Mechanism reproduces the confusion that may occur between objects in close proximity with each other (Craig et al., 2002). The process of forgetting is represented in ACT-R's base-level activation learning equation (Anderson and Lebiere, 1998).

Knowledge representation and model parameters are strongly constrained by the ACT-R theory and previous models (Craig et al., 2002; Anderson and Lebiere, 1998). However, individual differences and variations in strategies are inescapable realities of human cognitive behavior. The data collected in this Cognitive Probe project will be used to parameterize the ACT-R model. The prioritization distance errors can be modeled by setting the degree of similarity between object distance measures in ACT-R's Partial Matching Mechanism. The memory data can be modeled through the chunk retrieval threshold and retrieval noise parameters.

Constructive model runs will be conducted with scenarios of differing complexity, and the results compared to those obtained with the CART HPM's perfect implementation of the normative search strategy. While mission outcomes may not be notably different between the two HPMs (CART HPM vs. the TNM / ACT-R hybrid) for lower complexity scenarios, differences in shootlist management are expected to be observed (e.g., in terms of prioritization and memory errors). The CART HPM results correlated very well with the mission-level results using human pilots, but measures in that case study were not sensitive to variations in cognitive behavior. It is hypothesized that the number of vehicles in the CART scenarios did not push the limits of human performance, and that the pilots were able to compensate for their cognitive

limitations. Consequently, it is planned to conduct future experiments with additional levels of scenario complexity to determine the level at which the TNM / ACT-R hybrid model predicts CART HPM mission-level results should diverge from those observed in human-in-the-loop simulations. Model predictions will be tested with human subjects using the Cognitive Probe Testbed to determine the extent to which the TNM / ACT-R hybrid better represents actual human performance.

CONCLUSIONS

This project defined and implemented a concept for a virtual testbed for collecting data relevant to the cognitive task being modeled. During data collection, the pop-up probes employed in this testbed were found to be an effective means for capturing elements of cognition with minimal disruption. The same testbed will later be used to validate the models built using the data collected.

It was found that this cognitively-oriented testbed was not only important for providing model-parameterization data, but yielded significant insights regarding how well the testbed's virtual interface supported cognitive components of the SLM task. Even though it was not an objective of the Cognitive Probe study, shortcomings in the capabilities of the FRED interface to effectively support the SLM task were identified. The fact that the earlier CART HPM validation experiment reported by Martin et al. (2001) was insensitive to the cognitive aspects of the task caused researchers not to question the validity of the normative SLM search strategy. Cognitive Probe Testbeds could prove a highly useful adjunct to other modeling and simulation activities for defining and verifying effective tactics and interface configurations.

REFERENCES

Anderson, J.R., Bothell, D., Lebiere, C. & Matessa, M. (1998). *An Integrated Theory of List Memory*, Journal of Memory and Language, Vol. 38, pp. 341-380.

Anderson, J.R. & Lebiere, C. (1998). *The Atomic Components of Thought.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J.R., Matessa, M. & Lebiere, C. (1997). *ACT-R: A Theory of Higher Level Cognition and Its Relation to Visual Attention*, Human Computer Interaction, Vol. 12, pp. 439-462.

Boff, K. R., Kaufman, L. & Thomas, J.P. (1986). *Handbook of Perception and Performance*, Cognitive Processes and Performance, Vol. 2. New York, NY: John Wiley & Sons.

Brett, B. E., Doyal, J. A., Malek, D. A., Martin, E. A., Hoagland, D. G., & Anesgart, M. N. (2002). *The Combat Automation Requirements Testbed (CART) Task 5 Interim Report: Modeling a Strike Fighter Pilot Conducting a Time Critical Target Mission* (US Air Force Research Laboratory Technical Report AFRL-HE-WP-TR-2002-0018). Wright-Patterson Air Force Base, OH 45433-7022.

Craig, K. A., Doyal, J. A., Brett, B. E., Lebiere, C., Biefeld, E., & Martin, E. A. (2002). Development of a Hybrid Model of Tactical Fighter Pilot Behavior Using IMPRINT Task Network Modeling and the Adaptive Control of Thought – Rational (ACT-R). Proceedings *11th Conference on Computer-generated Forces and Behavior Representation*, Orlando FL, May 7-9, 2002. Simulation Interoperability Standards Organization, http://www.sisostds.org/.

DoD 5000.59-P. (1995). *Modeling and Simulation (M&S) Master Plan.* Defense Modeling & simulation Office, Alexandria VA 22311. http://www.dmso.mil

Gray, W.D. & Boehm-Davis, D.A. (2000). *Milliseconds Matter: An Introduction to Microstrategies and to Their Use in Describing and Predicting Interactive Behavior.* Journal of Experimental Psychology: Applied, 6(4), pp. 322-335.

Lee, F.J. & Anderson, J.R. (2001). *Does Learning a Complex Task Have to Be Complex?: A Study in Learning Decomposition*, Cognitive Psychology, Vol. 42, pp. 267-316.

Martin, E. A., Anesgart, M. N., Hoagland, D. G., & Brett, B. E. (2001). Goal-Oriented Human Performance Representation in DOD Constructive Simulations: Model Development and Validation. Proceedings *AIAA Modeling and Simulation Technologies Conference*, Montreal, Quebec, Canada, August 6-9, 2001. AIAA-2001-4415.

Miller, G. A. (1956). *The Magic Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information.* Psychological Review, 63, pp. 81-97.

Pew, R.W. & Mavor, A. S. (Eds.) (1998). *Modeling Human and Organizational Behavior: Application to Military Simulations.* Washington DC: National Academy Press.

Salvucci, D.D. & Anderson, J.R. (2001). *Integrating Analogical Mapping and General Problem Solving: The Path-mapping Theory*, Cognitive Science.

Schoelles, M.J. & Gray, W.D. (2000). *Argus Prime: Modeling Emergent Microstrategies in a Complex, Simulated Task Environment.* In N. Taatgen & J. Aasman (Eds.), Proceedings of the Third International Conference on Cognitive Modeling. Veenendal NL: Universal Press, pp. 260-270.

GLOSSARY OF ACRONYMS

ACT-R	Adaptive Control of Thought - Rational
AVC	Advanced Virtual Cockpit
CART	Combat Automation Requirements Testbed
COM	Microsoft's® Component Object Model protocols
DMS	Display Management Switch
DMSO	Defense Modeling and Simulation Office
DoD	Department of Defense
DOI	Display of Interest
EMC	External Model Call
FA	False Alarm
FRED	Fighter Requirements Evaluation Demonstrator
GCS	Generic Composite Scenario database
GMTI	Ground Moving Target Indicator
GPL	Global Prioritization List
GUI	Graphical User Interface
HITL	Human-In-The-Loop
HPM	Human Performance Model
HPMI	Human Performance Model Integration
ID	Object Identification Tag
JIMM	Joint Integrated Mission Model
M&S	Modeling and Simulation
NM	Nautical Miles
NRC	National Research Council
NTS	Next to Shoot
SAM	Surface-to-Air Missile
SGI	Silicon Graphics Incorporated
SIMAF	Air Force SIMulation Analysis Facility
SLM	Shootlist Management
Sub	Subject
TCT	Time Critical Target
TEL	Transporter / Erector / Launcher
TIR	Targeting Infrared
TMS	Target Management Switch
TNM	Task Network Model
TOI	Target of Interest
TSD	Tactical Situation Display
VSWE	Virtual Strike Warfare Environment

This page left blank intentionally

APPENDIX A
TNM / ACT-R INTERFACE

This page left blank intentionally

TASK NETWORK MODEL (TNM) / ACT-R INTERFACE

This appendix is a top-level summary of the TNM and ACT-R model interface requirements and ACT-R functional requirements. Since the intent was to obtain ACT-R shootlist management (SLM) model parameterization data from the testbed, this information was used as part of the planning process to determine features and data collection requirements appropriate for the Cognitive Probe Testbed. This information is included here to provide the reader insight in regard to the nature of the data to be passed between the TNM and the ACT-R SLM model, as well as some rationale for the testbed data collected.

Interface Requirements

Task Network Model Inputs to ACT-R

When developed, the ACT-R model will incorporate the shootlist prioritization criteria described in Table A-1. To enable ACT-R to perform the required shootlist prioritization task, ACT-R will need to be pre-populated with the aircraft's starting position and a population of representative

Table A-1. Shootlist Prioritization Criteria

Moving Status	Moving objects are given higher priority, as the scenario involves a mobile target that is likely to be on the move.
Range to the Reference Point†	Objects are prioritized based on their range from the reference point. Closer objects are given higher priority.
Time Since Failed Identification Attempt	Objects are temporarily assigned lower priority if the pilot attempts to identify them with the highest resolution sensor, and is unable to do so due to the aircraft's range from the object. Rather than dwelling on the object until it comes within identification range, the pilot will attempt to identify other objects on the list. The pilot will return to the particular object at a later time, when the aircraft is closer to it.

† Prior to the start of target acquisition activities, the shootlist contained only the 'Reference Point.' The 'Reference Point' was the latitude and longitude representing the best estimate of the target location. Since the target was a mover, the reference point position was stale by the time the pilot was trying to acquire the target.

error values from which it can randomly sample. These error values will be derived during the Cognitive Probe effort. Further, ACT-R must process data passed by the HPM during the course of a model run. These data include:

- Sensor used to take an image
- Number of objects in an image
- Data for each object in an image
    - Object type
    - Object size
    - Object latitude
    - Object longitude
    - Object elevation
    - Object range to the aircraft
    - Object range to the reference point
    - GMTI hit (yes or no)
    - Object moving (yes or no)
    - Object mover index
- Number of image objects detected or identified by HPM
- Number detected
- Number Identified
- Sensor used to take the image
- Data for detected or identified objects
    - Object ID
    - Object latitude
    - Object longitude
    - Object elevation
    - Object moving (yes or no)
    - Next sensor to be used for object
    - Detection result type (e.g., detected, detected and identified, etc)
    - Current simulation time
    - Time constant between repeated looks

- Aircraft latitude
- Aircraft longitude
- Reference point latitude
- Reference point longitude

ACT-R Outputs to HPM

Using the information above as well as specified prioritization criteria, the ACT-R model will perform the requisite calculations and then pass the following information back to the HPM. It should be noted that ACT-R will simply act as a repository for much of this information, passing unaltered data back to the HPM when needed.

- Sensor used to take an image
- Number of objects in an image
- Data for each object in an image
    - Object type
    - Object size
    - Object latitude
    - Object longitude
    - Object elevation
    - Object range to the aircraft
    - Object range to the reference point
    - GMTI hit (yes or no)
    - Object moving (yes or no)
    - Object mover index
    - Object previously detected (yes or no)
    - Object previously identified (yes or no)
- Number of items on the shootlist or 'priority list' (0-9)
- Number of items removed from shootlist
- Number of items added to shootlist

- Object of interest data

    Object ID

    Object latitude

    Object longitude

    Next sensor to be used for object

    Object range to the reference point

    Object range to the aircraft

    Object bearing to the aircraft

    Joint Integrated Mission Model (JIMM) object type

    Minimum time elapsed since last look achieved (yes/no)

    Point type (object or reference point)

ACT-R Functional Requirements

ACT-R Decision Making

ACT-R must use the inputs described above, coupled with specified prioritization rules, to determine the top nine priority objects. From this, it will determine answers to the following:

1. Which items get added / removed from the shootlist?

2. Which items get temporarily put on hold based on too great a range-to-target?

3. Of the items on the shootlist, which is the highest priority that is within sensor range and not currently on hold (i.e., the next to examine)?

ACT-R Representation of Human Information Processing Limitations

The ACT-R model will model cognitive performance to better represent human cognitive limitations. The ACT-R model that executes the shootlist management (SLM) task implements the same basic object prioritization strategy as the original CART model, however it incorporates the cognitive processes and effects that would potentially degrade operator performance to better represent human cognitive limitations. This degradation is represented in terms of sub-optimal prioritization (the pilot will not always perfectly implement prioritization rules) and forgetting

(the pilot may re-assign an object to the shootlist that he has previously identified as not being the target). The ACT-R implementation of this cognitive representation is summarized below.[11]

ACT-R implements a production system theory that models the steps of cognition through a sequence of production rules that fire to coordinate retrieval of information from the environment and from memory. It is a cognitive architecture that can be used to model a wide range of human cognition. It has been used to model tasks as simple as memory retrieval and visual search to tasks as complex as learning physics and designing psychology experiments. *Goals are a central concept in ACT-R that correspond directly to tasks in a CART TNM.* A goal in ACT-R is a declarative structure that encodes a particular objective (e.g., perform a sequence of actions, or find an answer to a question) that is the current focus of attention. Each production rule in an ACT-R model applies to a specific type of goal. When a goal is solved, it is stored in declarative memory as a structure (called a chunk) that encodes the result of that goal. Thus a type of goal, together with the production rules that apply to it and associated declarative chunks, can be thought of as a modular piece of knowledge. Models of complex cognitive tasks can be built around the assembly of multiple goals.

The first ACT-R goal corresponds to the SLM task to update the Objects Of Interest (OOI) list. That list in ACT-R typically held six or fewer objects, since memory chunks in ACT-R -- as with humans -- are constrained to hold only a small, fixed number of items. A set of memory chunks is created that encode -- for each target -- its basic characteristics (ID, moving status, latitude and longitude) and whether it has been detected and / or identified. The CART task network passes that basic target information to the ACT-R model whenever this goal is called.

The second ACT-R goal is to filter the Image List resulting from the processing of the radar screen by the task network. For each target, given its description (ID, latitude, longitude), ACT-R attempts to retrieve from memory whether that target has been previously detected and / or identified. To do that, ACT-R simply attempts to retrieve from memory a chunk created by the goal to update the OOI list that states that the target has been detected or identified. If the retrieval fails, then the model assumes that the target has not been detected or identified.

[11] A more detailed treatment may be found in Craig et al. (2002).

However, memory retrievals in ACT-R -- like human memory -- are far from perfect. Through ACT-R sub-symbolic level processing, it is possible that the retrieval of an object that has been encoded in memory as 'identified' might fail. As a result, the model might decide to examine that target again. Moreover -- unlike other production systems in which matching chunks in memory to production conditions is a perfect process -- in ACT-R all chunks of the same type compete for any given retrieval, with chunks that only partially match the desired pattern having their activation penalized by an amount that reflects the difference between pattern and chunk. This partial matching mechanism in ACT-R reproduces the confusion that may occur between objects in close proximity to each other. In this case, even though a target might not have been previously seen, if another close-by target had been detected and / or identified, it might lead to an erroneous classification of the original target, which is then omitted from the search. Thus probabilistic retrieval from memory can lead to occasional errors in which a target is examined multiple times or not at all.

The third goal corresponds to the prioritization of the shootlist. The task starts by recalling the location of the reference point. It then implements the original CART prioritization rule by attempting to retrieve the closest moving target to that point that has not already been selected. Since this is a probabilistic partial matching process, the rule is only approximately implemented; targets slightly further from the reference point could be selected ahead of closer ones. Moreover, an object is selected only if there is no prior memory of it being identified, which (as was discussed for the previous goal) can lead to both omitted and repeated identifications. Finally, after an object is selected, its position becomes the current focus of attention around which the search for the next target will start. Thus, while the process still favors targets closest to the reference point, a tendency toward selecting targets in clusters arises. This is compatible with the memory requirements of the task, since remembering that a cluster of targets has been detected and identified is much easier than remembering the same number of scattered points.

APPENDIX B
SCENARIOS

This page left blank intentionally

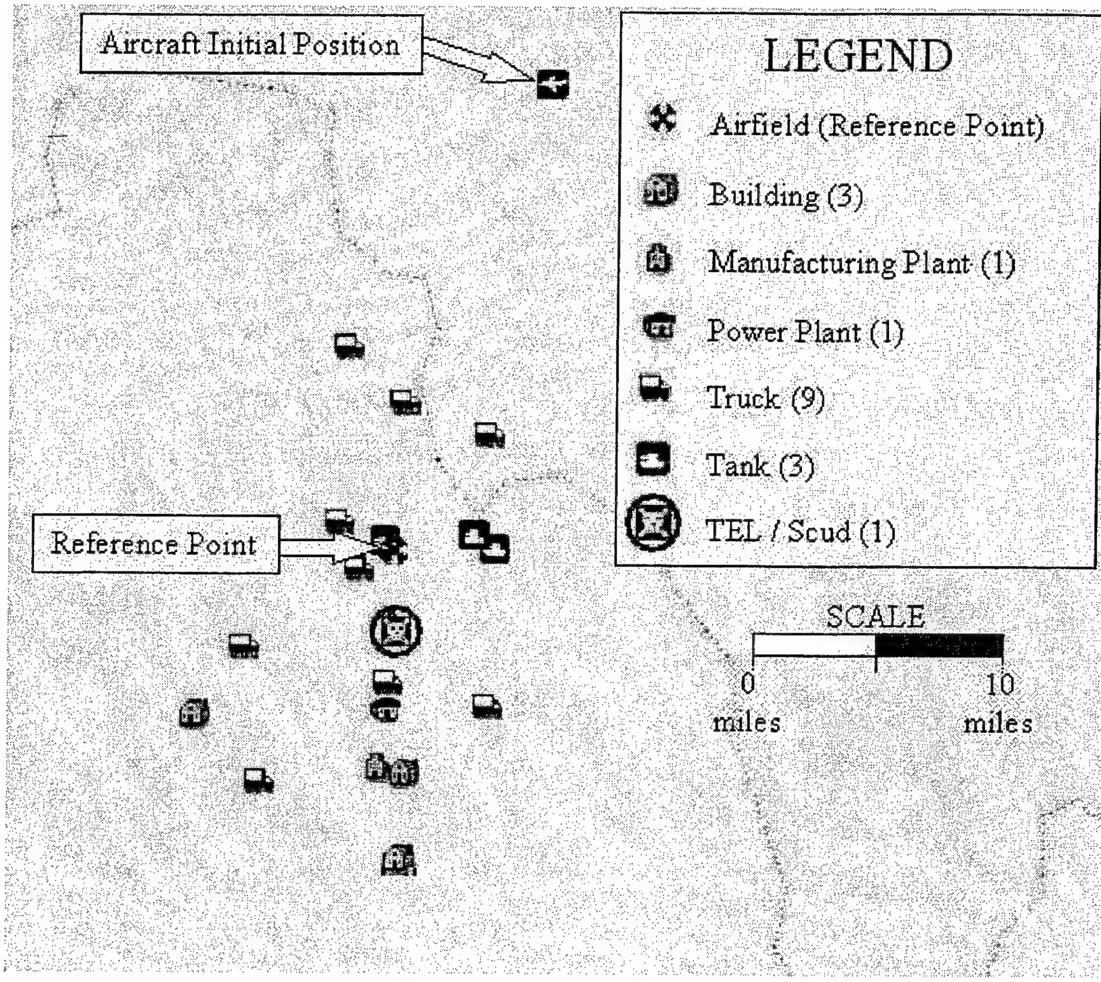


Figure B-1. Practice Scenario.

Figure B-1 shows the initial condition for the lay down of buildings and movers used for the Practice Scenario. The numbers in parentheses in the legend indicate the quantity of each entity type placed in the gaming area. The Reference Point represented the information provided the subjects as the best estimate of the Scud Transporter / Erector / Launcher (TEL) vehicle location. The Reference Point was the airfield control tower.

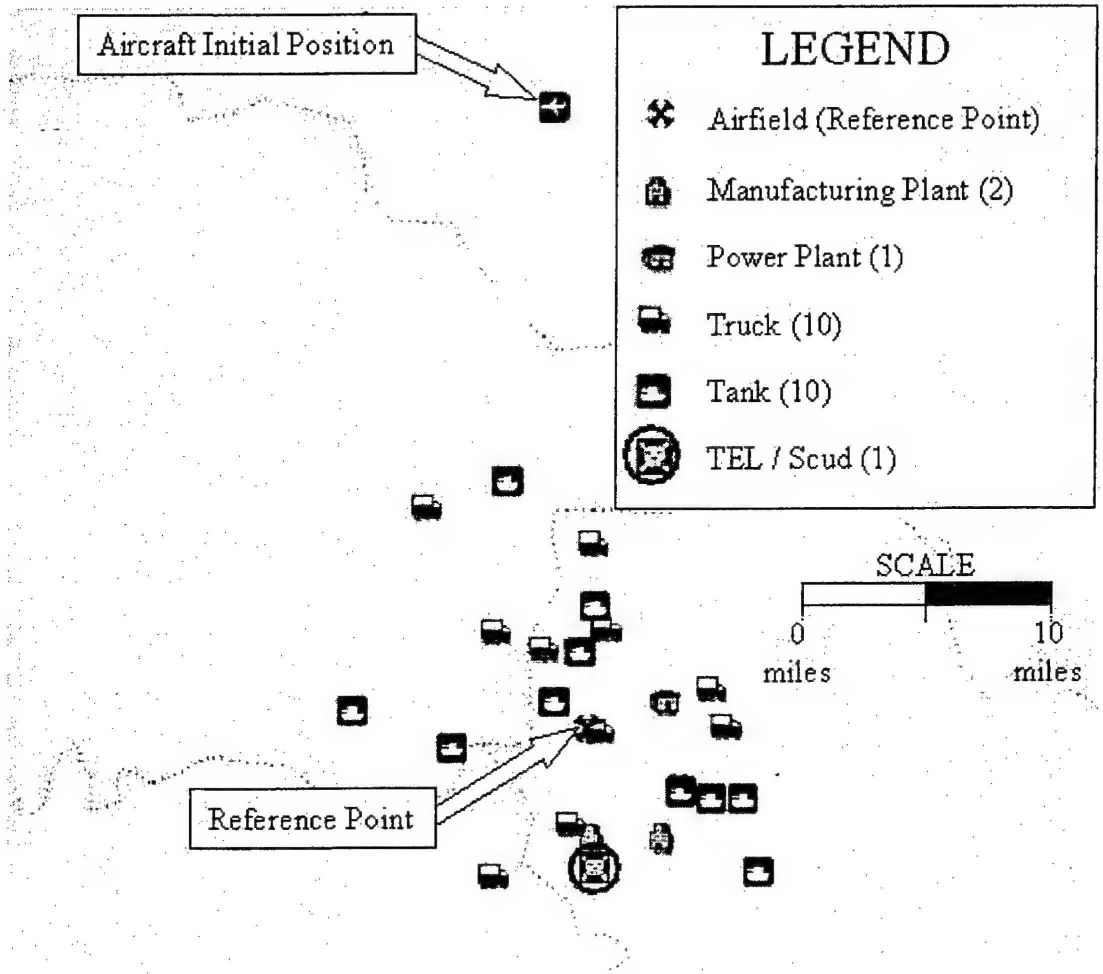


Figure B-2. Data Collection Scenario 1.

Figure B-2 shows the initial condition for the lay down of buildings and movers used for Data Collection Scenario 1. The numbers in parentheses in the legend indicate the quantity of each entity type placed in the gaming area. The Reference Point represented the information provided the subjects as the best estimate of the Scud Transporter / Erector / Launcher (TEL) vehicle location. The Reference Point was the airfield control tower. The TEL was on the far side of the Reference Point in this scenario so that it wouldn't appear in the subject's initial shootlist, thus requiring the subject to remove and add objects from the shootlist after entering TIR range; this provided a better opportunity for obtaining memory data than did Scenario 2.

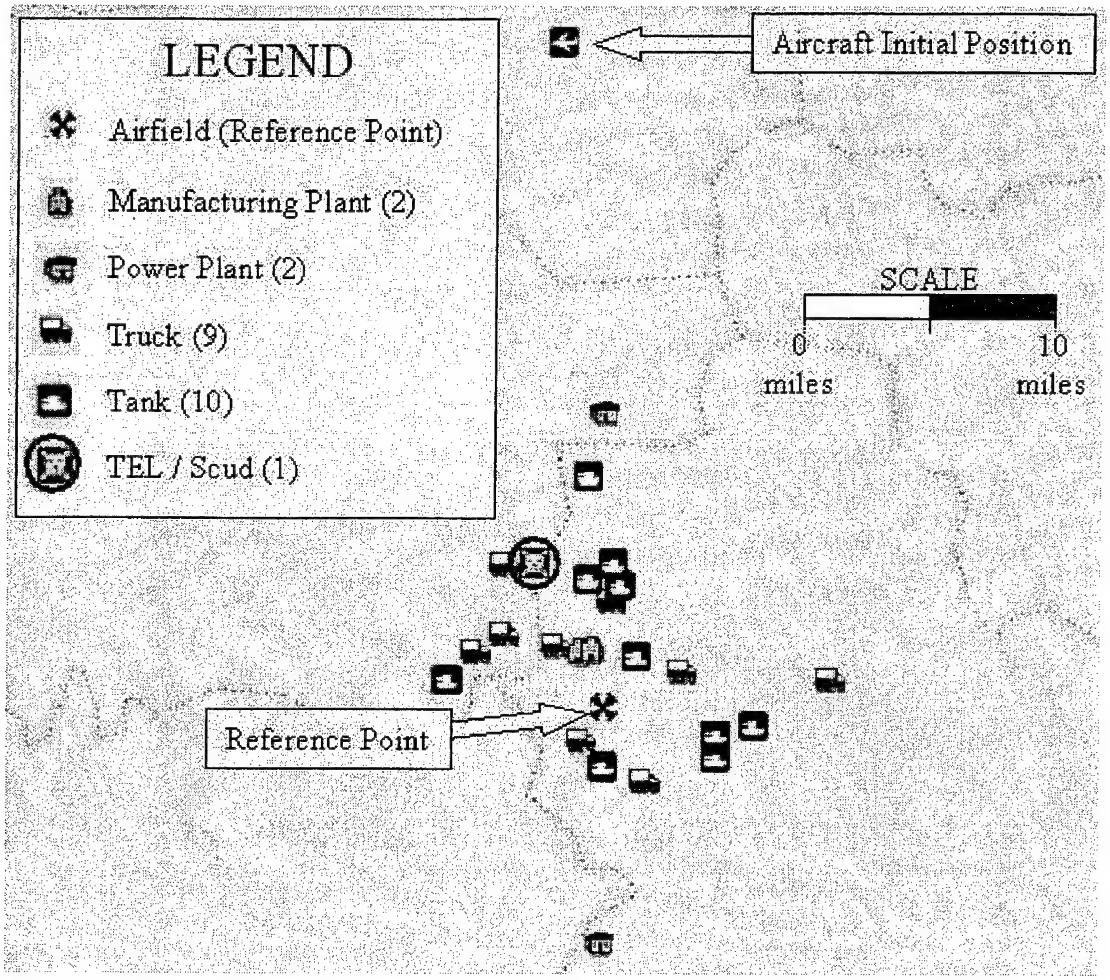


Figure B-3. Data collection scenario 2.

Figure B-3 shows the initial condition for the lay down of buildings and movers used for Data Collection Scenario 2. The numbers in parentheses in the legend indicate the quantity of each entity type placed in the gaming area. The Reference Point represented the information provided the subjects as the best estimate of the Scud Transporter / Erector / Launcher (TEL) vehicle location. The Reference Point was the airfield control tower.

This page left blank intentionally

APPENDIX C
INSTRUCTIONS TO THE SUBJECTS

This page left blank intentionally

General information

The probe procedure is divided into three stages. The first stage is training followed by practice and then data collection. During training, you will be instructed on how to interact with the interface. Also, the details of the part-task TCT mission and a suggested search strategy will be discussed. This will be followed by the practice stage. Here you will familiarize yourself with the controls and ask questions to clarify any ambiguities. A checklist of critical tasks (see Table C-1) will be completed to ensure that you have sufficient proficiency to complete the mission.

Table C-1. List of Critical Tasks Subjects Must Master Prior to Data Collection

CRITICAL TASKS CHECKLIST	
Change display of interest	
Add item to shootlist on GMTI	
Add item to shootlist on TIR	
Remove an (1) item from shootlist	
Remove all items from shootlist	
Deploy TIR	
Zoom in on GMTI	
Zoom in on TIR	
Inspect (scroll through) items on shootlist	

Following completion of the practice run(s), the first of the two data collection runs will begin. Mission time for the two runs will be approximately 15 minutes. Including training, practice, and between-scenario downtime, total time of your participation is estimated at less than two hours.

Mission

For each trial your mission is to find and destroy a time critical target (TCT). The TCT is a mobile scud transporter erector launcher (TEL). To aid in this process you will be given a

reference point indicating the last known location of the TCT. The reference point latitude and longitude is unique for each scenario; however, this point is located at the control tower of the airport in all scenarios.

You will search for the moving Scud TEL using onboard sensors to develop a "shootlist". The shootlist is a number of icons on the tactical situation display, radar display, or targeting infrared (TIR) display that highlight the points/objects of interest that you designate. When you designate an item to the shootlist a white broken triangle appears around it. The current object of interest or "next-to-shoot" on the shootlist is indicated by a solid triangle. The shootlist is built and modified dynamically over the course of the mission. You may add/delete items from the shootlist, as you desire; however, the maximum number of shootlist slots is nine. If you exceed this, the first item added to the list is automatically removed or "bumped" from the shootlist. You may use any of the displays to designate an object or remove ALL objects from the shootlist; however, the simulation software has a "quirk" that only allows single-item removal from the TIR display. If you wish to remove a single item from the shootlist, select the TIR display and then remove the object.

Sensors

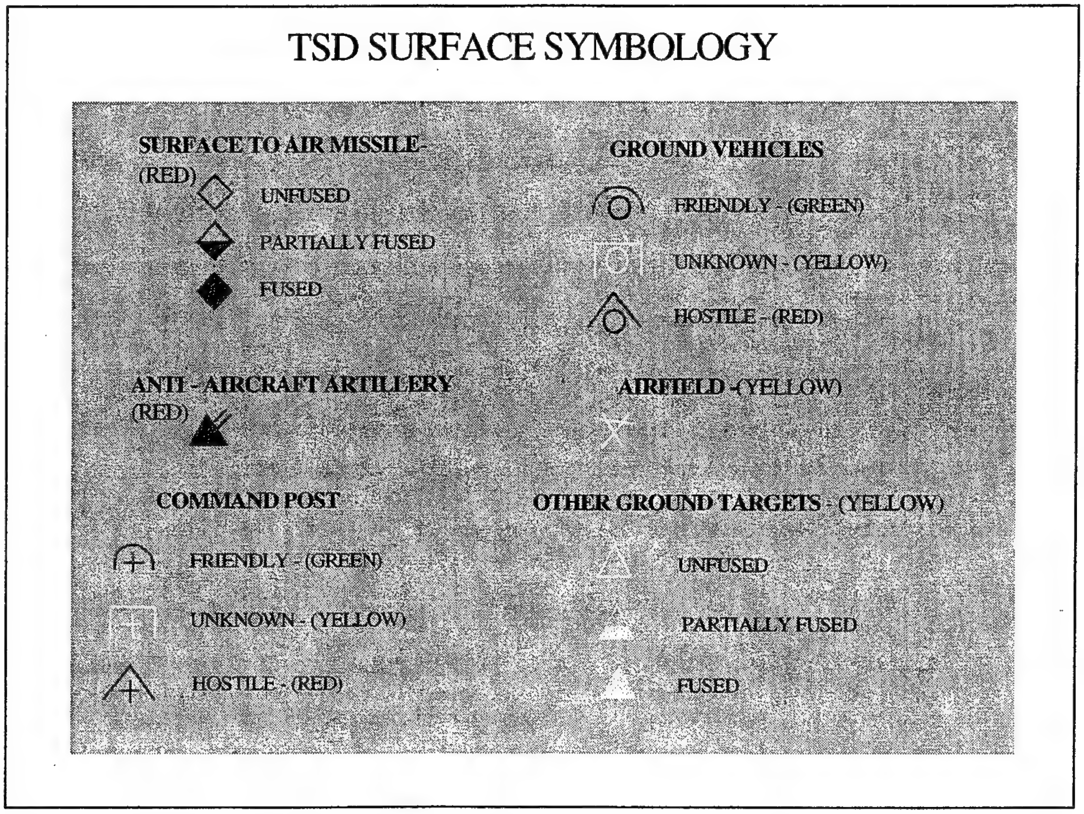
Three sensor displays are available to search and detect the TCT. The left panel is the real beam radar with a GMTI overlay, the middle is the tactical situation display (TSD), and the right panel is the targeting infra-red (TIR) display. These displays are described in more detail below.

GMTI
Ground Moving Target Indicator
This display is used to DETECT objects. Yellow dots indicate a moving object. Green "eyebrows" indicate a stationary object or a mobile object that the radar has not detected as moving. To zoom out on this display, move the cursor up to the top of the panel. To zoom in, move the cursor down.

TSD

Tactical Situation Display

This display is used to DETECT objects. The TSD displays the aircraft's waypoints along with GMTI hits and other ground objects. The symbols used on the TSD are in the figure below. To zoom out on this display, move the cursor up to the top of the panel. To zoom in, move the cursor down.



TIR

Targeting Infra-Red

This display is used to IDENTIFY objects. The TIR sensor must be manually deployed. This is accomplished by pressing the bottom left button on the fly panel. There are three different field-of-view (FOV)/ zoom settings available for the TIR sensor (wide, narrow, and 2X narrow). The bottom right button on the fly panel can be used to cycle through the three levels. The current FOV or zoom level is indicated on the left side of the TIR display.

Tactic

You initially detect moving targets using the GMTI capability on the radar display or on the TSD. Designate these detections to add them to a shootlist by moving your cursor to the object and pressing the middle switch upwards. **Choose the nine closest detections to the reference point**, as these are the nine objects most likely to be the target. Once you sequence waypoint 2, you should be within TIR range (approximately 20 NM) of the target area. Do *not* deploy the TIR until you reach this waypoint, as it provides little value at greater ranges and it also increases your radar cross section, making you more vulnerable to air to ground threats. (Although, such threats will not be played out in this scenario). Once you have populated your shootlist and have sequenced waypoint 2, deploy your TIR and begin trying to identify objects on your shootlist. Because there are more movers on the ground than you have slots in the shootlist, you may have to periodically remove one/all shootlist object(s) to make room for a different object. Once you have positively identified the target, make that target the NTS, and pull the trigger switch on the joystick. At this point the experimenter will end the trial. If you do not succeed in finding the target, the experimenter will end the trial once the target area is no longer in TIR range.

Controls

Display management switch (DMS) – The DMS is the right-most button on the joystick, and it is used to control the Display of Interest (DOI). Moving the switch in any of three directions selects the corresponding display as the DOI. Moving the switch *right* selects the TIR, moving the switch *down* selects the TSD, and moving the switch *left* selects the radar with GMTI. The DOI is designated on the display by green highlighting around the perimeter.

Target management switch (TMS) – The TMS is the center button on the joystick. It permits the designation, rejection, and selection of target locations on multiple displays. The TMS functions described affect the TSD, radar, and TIR displays. Moving the TMS to the *forward* position designates the current cursor position. When a target location is designated it is also added to the shootlist. Moving the TMS *forward* again will make that target the current target of interest denoted by a solid triangle (also known as the Next To Shoot – NTS). The shootlist will hold up to 9 targets. To step through the shootlist move the TMS *left*. Subsequent movements

will continue to cycle through the shootlist. To undesignate (or reject) the current NTS-target, move the TMS *down and release in less than 1 second.* To reject a target that is not the NTS, capture the target by the cursor (cursor over the target) and move the TMS *down and release in less than 1 second.* To clear the entire shootlist (remove all designations) move the switch *down and hold for more than 1 second.*

Trigger -- The trigger is used to release an air to surface missile on the target. When you have identified the scud depress this button. Once the weapon is released the mission is considered complete.

Fly Panel – A series of buttons appear on the left of the joystick. Only the bottom two orange colored buttons will be used.

Deploy TIR -- The lower-left button controls deployment of the TIR. When the TIR is retracted pressing, the button will deploy the TIR. Subsequent presses will alternate between deployed and retracted positions for the TIR.

Zoom TIR – The lower-right button controls the field-of-view (FOV) or zoom level of the TIR. The three levels are wide, narrow, and 2X narrow. Depress the button to move through the levels starting at wide moving to narrow then to narrow 2x and then back to wide.

On-line Survey

When you remove or bump an item(s) from the shootlist, the simulation will pause and a button will appear on the upper left corner of the interface. Click this button to reveal a drop down menu containing five choices. These choices are explained in detail below. Choose the selection that best describes your rationale for removing the item.

Removed-Identified

This option is intended to represent instances where you have identified the object using the TIR display as *not* the TCT and have purposely removed it. In other words, you have identified the object of interest as either a semi or a tank.

Removed- Unknown

Choose this option when you have intentionally removed an object but do not know what it is. An example of when you would want to select this option is if you want to add an additional object to the shootlist but your shootlist is full; you may choose to remove an unidentified object that is further from the reference point to make room for this item.

Removed – Mistake

Choose this option if you remove a designation to the shootlist that you either added by mistake or removed by mistake(e.g. hit wrong switch).

Bumped – Identified

Select this if you have identified the object and it is bumped from your shootlist. This could occur if you identified all objects on your shootlist then added a tenth item.

Bumped- Unknown

Select this if a item is bumped from your shootlist and you have not identified it.

APPENDIX D
SUBJECTIVE SEARCH STRATEGY DATA

This page left blank intentionally

A questionnaire was administered to help identify the subjects' strategies for building the shootlist. This was intended to provide insights regarding potential modifications necessary for the modeled search strategy. The questionnaire was given at the completion of the two data collection trials. Subjects were given no prior knowledge of the two questions in order to preclude such knowledge from influencing their behavior. *'**Subject Responses**' are sequenced in the same subject order for both questions.*

Question 1: *During the experiment, did you implement the shootlist strategy discussed in the instructions?*

Subject Responses to Question 1:

- Yes, I looked at the items closest to the reference point and followed suggestion of looking at others when unable to identify the first selected items.
- Yes
- For the most part, yes I did. The variation was the fact that I slightly biased the closest nine objects to the reference point to those closest to the planned aircraft approach. The theory was that I could identify them sooner and then continue the search of objects that were on the "other side" as I got closer to the reference point.
- Yes, I followed the strategy as briefed.
- Yep!
- I did for the most part. Initially, I would select a set of targets around the reference point and try to leave one shoot list slot open. I intended to use this open slot to designate and look at targets on the way into the target area. While it seemed like a good idea, I tended to get confused in the remove process and didn't always remove the target in the open slot but removed another target near the reference point. This led to confusion about whom I had designated in the target area.

Question 2: *Given what you experienced in the experiment, explain the strategy you would use to manage the shootlist to search and identify the TCT. (Consider your strategy in a real-world context that includes terrain, control of the aircraft, potential of threats, many more movers, more types of movers, etc.)*

Subject Responses to Question 2:

- To look at more items sooner in order to eliminate potential targets even those not close to the reference point.

- I think I would probably add the targets to the shootlist that were surrounding the last known position of the target of interest and wait until I was closer to the targets for quicker identification.

- I would probably divide the area around the reference point into sectors and look at the movers / objects in the area closest to the ingress aircraft position first.

- Given the relatively few movers (20 or less) and the relatively long time I had to acquire (which is a function of airspeed and TIR range), I probably would have had time to examine the targets closest to the aircraft's ingress path first. Instead, I examined those closest to the reference point first. My suggestion is that if there are few targets and plenty of time to examine them, examine the ones closest to the aircraft first (before you overfly them). If there are a larger number of targets, or if acquisition time is more limited due to a higher airspeed or shorter sensor range, I think the strategy briefed is a good one.

- Elected to use GMTI first to identify targets at around the 30 mile range point. Tried to nominate as many targets as I could between 30 and 20 miles using GMTI. At least for the first nine. Transitioned to TIR at 20 miles range. Scrolled to first target nominated. Tended to try and ID the targets when they were between 20 and 16 miles. On several occasions I could find certain trucks and tanks immediately. Most of the time, however, I made decision to remove targets on the list between 15 and 11 miles range. In case you haven't noticed, I used range to cursor a lot. Tried to get

through the first nine before going on to second nine. On one occasion, I elected to leave two targets until I got under 10 miles range to cursor for positive ID. While I was waiting to get closer, I went out and nominated targets to fill up the shoot list. When I did that, I checked the radar display for range to GMTI hits then transitioned to the TSD for GMTI hits selecting only the staple and circle. Once I found the TEL, I killed it then ejected.

- I think I would use the same basic strategy described above *(i.e., in response to Question 1)* with some slight modifications. I would still designate an initial set of targets near the reference point but would focus on targets on my side of the reference point. I would still leave a shoot list slot open to evaluate targets as I come across them (particularly movers). Once in the target area I would work the targets I had designated initially. If I don't find the TCT, I will designate targets on the other side of the reference point and evaluate them. I would continue this process as I work away from the reference point until I find the target. If there were multiple airplanes I would divide potential targets between them (e.g., one takes targets on the right side, one takes targets on the left). If target areas are obscured by terrain, I would split the flight to approach from different directions so we can cover the target area completely.